ANATOMY AND PROFESSIONALISM IN AN UNDERGRADUATE MEDICAL CURRICULUM

FINN, GABRIELLE,MARIA

ANATOMY AND PROFESSIONALISM IN AN UNDERGRADUATE MEDICAL CURRICULUM

Gabrielle Maria Finn

School of Medicine and Health

Durham University, UK

2010

Doctor of Philosophy (PhD)

Abstract

Name: Gabrielle Maria Finn

Title of thesis: Anatomy and professionalism in an undergraduate medical curriculum

Higher degree for which submitted: Doctor of Philosophy (PhD)

Year of submission: 2010

This thesis describes two themes within the undergraduate medical curriculum; innovations in anatomy teaching, and the assessment of professionalism. Methodologies from both the quantitative and qualitative paradigms were utilised.

The main findings were:

1) The Virtual Human Dissector™ (VHD) was shown to be equally as effective as cadaveric prosections as a tool for learning cross-sectional anatomy.

2) Body painting was demonstrated as being a highly motivating and engaging exercise for students. Students reported that the bold colours and kinaesthetic nature of body painting promotes retention of knowledge and informed their approach to future patients when painting was coupled with simultaneous peer-physical examination (PPE).

3) Contextual learning and simulation were shown to directly impact upon retention of knowledge through the use of clothing in anatomy education. This highlighted how when implementing simulation small and seemingly trivial details, such as clothing, are important.

4) The Conscientiousness Index (CI) has been demonstrated as an objective and scalar measure of one element of professionalism, conscientiousness. The CI identified students at the positive and negative end of the behavioural spectrum, and this correlated with peer and staff judgements on the professionalism exhibited by students at these extremes of behaviour.

5) Students were able to accurately assess the conscientiousness of their peers, however were unable to self-assess conscientiousness. The reliability of

such peer assessments was improved when peers assessed only those in their tutor groups, with whom they had the majority of academic contact, compared to when assessing the entire cohort. This demonstrated the importance of assessor familiarity in assessments.

6) Critical incident reporting, of extremes in professionalism, was shown to promote reflection in students. Critical incident reports, as with the Conscientiousness Index, offers faculty a tool by which outlying students can be identified.

7) Students were unclear about the constituent elements of professionalism and the contexts in which professionalism was relevant. Three contexts were identified; the clinical, the academic (University), and the virtual (online) context. The impact of professionalism assessments and the scrutiny on students has led students to struggle with identity negotiation. This was with respect to their personal and professional identities and the expectations of different stake holders, such as faculty, the media and prospective patients.

Table of contents

List of tables

List of figures

Declaration

The material contained in the thesis has not previously been submitted for a degree in Durham University or any other institution.

Chapters 3, 6, 7 and 8 contain joint research. The contributions of individual authors are described within each Chapter.

Statement of Copyright

Acknowledgments

I would like to thank my supervisor Professor John McLachlan for all of his time, effort and enthusiasm over the last 3 years. He has patiently answered all of my many queries and has taught me endless amounts about research.

Dr Debra Patten for her support in implementing my research within the anatomy curriculum and for her efforts as a supervisor.

Dr Marina Sawdon for all of her hard work, support and friendship. Without Marina, a lot of the data collection and analysis would not have been possible.

Dr Fiona Curtis for being my e-learning guru.

Thank you to all the medical students at Durham University who have participated in my research projects, I am extremely grateful that you all indulged me. Thank you to all the staff at Durham University Phase 1 Medicine for their assistance in all aspects of my research.

Publications from this thesis

- Donnelly JL, Patten D, White P and **Finn, G** Virtual Human Dissector as a learning tool for studying cross-sectional anatomy. Medical Teacher (2009). 31(6):553-555.

- McLachlan JC, **Finn GM,** and Macnaughton RJ. The Conscientiousness Index: An objective scalar measure of conscientiousness correlates to staff expert judgements on students' professionalism. Academic Medicine (May 2009). 84(5): 559-565

- Hodges D, McLachlan, JC and **Finn GM.** Exploring reflective 'critical incident' documentation of professionalism lapses in a medical undergraduate setting. BMC Medical Education. (2009). 9:44

- **Finn GM,** Sawdon MA, Clipsham L and McLachlan JC. Peer estimates of low professionalism correlate with low Conscientiousness Index scores. Medical Education. (2009).43: 960-967.

- **Finn GM,** Patten D and McLachlan JC. The Impact of wearing scrubs on contextual learning. Medical Teacher (2010) 32: 381–384.

- **Finn GM** and McLachlan JC. A qualitative study of students' perceptions of body painting. Anatomical Sciences Education (2010) 3:33-38.

- **Finn GM,** Garner J and Sawdon MA. "You're judged all the time!" Students' views on professionalism: A multi-centre study. Medical Education (2010). 44: 814-825.

- **Finn, GM**. Twelve tips for running a successful body painting teaching session. Medical Teacher (in press).

Papers in review

- **Finn GM** and Garner, J. Twelve tips for conducting a successful peer assessment exercise. Medical Teacher.

- Tiffin PA, **Finn GM** and McLachlan, JC. Can Professionalism be Assessed through Multiple Choice Questions? BMJ.

Conference contributions from this thesis

- McLachlan JC, Patten D, **Finn GM**, White P, Richardson S. Living Anatomy in the undergraduate medical curriculum – new ideas (Body painting, Ultrasound, Peer-physical examination). Annual meeting of the Anatomical Society of Great Britain and Ireland. Durham, July 2007.

- Donnelly JL, Patten D, White P, **Finn GM**. Virtual Human Dissector as a learning tool for studying cross-sectional anatomy. AMEE (The Association for Medical Education in Europe) conference in Trondheim, Norway. August 2007.

- McLachlan JC, Finn GM and Macnaughton RJ. The Conscientiousness Index: An objective scalar measure of conscientiousness correlates to staff expert judgements on students' professionalism. Invited presentation: ASME Conference, 5th June 2008, London (Medical Students and Professional Behaviour)

- **Finn, GM,** Patten, D and McLachlan JC. The Impact of wearing scrubs on contextual learning. Poster, ASME Annual Conference New Horizons in Medical Education. 10-12th September 2008.

- McLachlan JC, Sawdon MA, Macnuaghton RJ and **Finn GM**. Assessment of conscientiousness and its relation to professionalism. Presentation: Teaching and Learning Conference: Improving the Experience, 4th September 2008, Newcastle University, Faculty of Medical Sciences. (**Prize winning presentation and abstract**)

- **Finn, GM.** Making Use of the Living: Innovative Methods of Teaching Living Anatomy. Poster. Durham University Doctoral Fellowship Event. 19th February 2008. (**Prize winner: Best poster as voted for by the entrants**)

- **Finn, GM.** Making Use of the Living: Innovative Methods of Teaching Living Anatomy. Poster. Yorkshire and North East UKGRAD Poster Competition and Networking Event, University of York. 7th March 2008.

- **Finn, GM**.The Colour of Your Inside: Using the Body as a Canvas. (Invited key speaker). "Create" conference on colour. University of Ulster, Northern Ireland. October 2008.

- **Finn, GM.** The Art of Anatomy: Alternative Approaches to Teaching Anatomy. Winter Meeting of the Anatomical Society of Great Britain and Ireland. Oxford University. 6-8th January 2009.

- **Finn, GM.** The Impact of Scrubs on Contextual Anatomy Learning. Poster. Experimental Biology (American Association of Anatomists). New Orleans 2009. April 16-19th 2009.

- **Finn, GM.** The Colour of Your Inside: Anatomical Body Painting. Presentation. Experimental Biology (American Association of Anatomists). New Orleans 2009. April 16-19th 2009.

- **Finn, GM** and Sawdon, MA. Does peer and self-assessment correlate to the use of the Conscientiousness Index when evaluating professionalism in medical students? Poster. ASME Annual Scientific Meeting. The Royal College of Physicians of Edinburgh 15-17 July 2009.

- **Finn, GM** and Sawdon, MA. Students' views on peer assessment and professionalism: knowing when to "switch it on." Poster. ASME Annual Scientific Meeting. The Royal College of Physicians of Edinburgh 15-17 July 2009.

- McLachlan JC, **Finn GM**, Sawdon M, Macnaughton J, Clipsham L, Douglass S. Exploring the relationship between professionalism and conscientiousness. Full Research Paper. AMEE International Conference Malaga September 2009.

- Fleming K, McLachlan JC, **Finn GM**, Ludlow A. Invited Speakers Flex and Ply: does my S3 look big in this? AMEE International Conference Malaga September 2009.
- **Finn, GM** and Sawdon, MA. Students' views on peer assessment and professionalism: knowing when to "switch it on." Presentation. AMEE International Conference Malaga September 2009.
- **Finn, GM** and Sawdon, MA. The impact of feedback and personality upon students' ability to self-assess skills, attributes and aptitudes. Ottawa Conference, Miami May 2010.Presentation.
- Fleming K, **Finn GM** and McLachlan JC. Does my S3 look big in this? Crafting invisible maps of the body. Ottawa Conference, Miami May 2010.Presentation.
- **Finn, GM** and Sawdon, MA. The impact of feedback and personality type upon students' ability to self-assess skills, attributes and aptitudes. ASME July 21-23 2010. Cambridge, UK.
- McLachlan JC, **Finn GM** and Sawdon MA. The relationship between conscientiousness and professionalism. ASME, 21-23 July 2010, Cambridge.

- McLachlan JC, **Finn GM** and Sawdon MA. “With all due diligence”: measuring conscientiousness in complex settings. Ottawa Conference, Miami. May 2010.Workshop.

- **Finn, GM**. CETL4HealthNE: Improving the future healthcare workforce conference .19 March 2010, Newcastle. The use of Lectopia™ in medical education. Poster.

- **Finn, GM**. CETL4HealthNE: Improving the future healthcare workforce conference .19 March 2010, Newcastle. Dr Companion and the use of mobile technologies to support student/staff access to electronic teaching materials. Presentation.

- Moss, J, Scott L and **Finn, GM**.CETL4HealthNE: Improving the future healthcare workforce conference .19 March 2010, Newcastle. Learning technologies to support student/staff access to electronic teaching materials. Workshop.

- **Finn, GM**. Using art in medical school. Researchers Revealed. Hancock Museum 2009.Presentation and exhibition.

- **Finn, GM** and McLachlan JC. Designer Bodies. Life Centre, Newcastle. 18th March 2010. Presentation, workshop and exhibition.

- McLachlan, JC and **Finn, GM.** The Appliance of Fashion to Science. Durham County Hall, 25th March 2010. Presentation.

- **Finn, GM** and Curtis, FA. The art of anatomy teaching**.** 3Rivers Consortium conference which is taking place on 30th March 2010 at Sunderland University.

- Fleming, K, McLachlan, JC and **Finn GM**. Invisible maps of the body: constructing myths, uncovering legends. 4th International Conference of the European Society for the History of Science. Barcelona, 18-20 November 2010.

- **Finn GM** and McLachlan JC. Students' perceptions of body painting as a tool for learning anatomy. AMEE conference, Glasgow, 4-8 September 2010. Presentation.

- **Finn GM**, Sawdon MA and Garner J. Students' attitudes towards peer assessment of professionalism: a multi-centre study. AMEE conference, Glasgow, 4-8 September 2010. Presentation.

- **Finn GM**, Curtis FA, Cox DR and Northend M. The use of videos to aid self-directed learning and improve understanding of living anatomy. AMEE conference, Glasgow, 4-8 September 2010. Presentation.

- **Finn GM**. The Art of Anatomy: using colour in medical education. Create, Conference on Colour. Gjovik, Norway. June 2010. **Keynote lecture**.

- **Finn, GM.** Colourful colons: using colour in medical education. “Working Together, Learning Together” Conference. James Cook University Hospital, Middlesbrough. Monday 12th June. **Keynote lecture.**

Abbreviations

ABIM= American Board of Internal Medicine

AIDS = Acquired Immune Deficiency Syndrome

BP= Body paint

CI= Conscientiousness Index

CT = Computed Tomography

DR = Dissecting Room

GMC = General Medical Council

GP= General Practioner

LD= Line drawing

MRI = Magnetic Resonance Imaging

MSF= Multi-source feedback

NHS= National Health Service

OBC= Objectified Body Consciousness

PBL = Problem- Based Learning

PPE= Peer physical examination

SDL= Self Directed Learning

UDQC= University of Durham, Queen's Campus

UK= United Kingdom

USA= United States of America

VHD= Virtual Human Dissector

VLE= Virtual Learning Environment

Chapter 1: Introduction

1.1 Introduction to the introduction

This thesis, titled “Innovations in anatomy and professionalism in an undergraduate medical curriculum”, reports on two themes; anatomy and professionalism.

The original theme for this thesis had been innovations in anatomy pedagogy. I had planned to measure the efficacy of innovative pedagogic approaches and relate these to principles of education, specifically medical education. Despite achieving this goal, professionalism became another focus for this thesis.

Research is opportunistic. This is certainly true for the research undertaken during my doctoral studies. During the first year of my studies Professor John McLachlan initiated the Conscientiousness Index, described in later chapters. I was fortunate to be a part of this project, which related to medical professionalism, and found the direction of this preliminary research to be of great interest. The initial study generated interesting findings and it became evident that this is where the important research lay.

Subsequently, the Medical Education Research Group (MERG) was born. This group served as a forum for generating and discussing research ideas. As new opportunities arose within MERG, predominantly relating to measuring professionalism within our undergraduate students, I perceived the opportunity to make a more significant contribution to the literature by exploring professionalism.

Although anatomy and professionalism are seemingly distinct themes, these two strands have proceeded in parallel. As my studies have progressed these themes have added to and enriched each other.

As will be described within the final discussion (chapter 10), common themes have emerged from both areas of research. These include identity and the role of clothing in education. I feel that conducting research in two parallel subjects, both relating to medical education, has provided me with a better understanding of the principles of medical education and has enriched this thesis.

This introduction will review the literature for both anatomical and professionalism education in turn.

1.2 Anatomy

Anatomy may no longer be considered a discipline in which there are many scientific advances [1], but it is at the heart of medical education research. In order to fully understand the developments being made today we need to understand from where anatomy has come.

This part of the introduction will provide an insight into anatomy and its teaching, from its origins with Galen and Vesalius to the modern day and its recent explosion into the spotlight after publicity surrounding the German anatomist Gunther von Hagens.

There has been ongoing debate as to the best way to teach anatomy for many decades, whether it be by dissection, prosection or the use of new technologies[2-7]. In this introduction I hope to summarise the previous debates and offer my views on the topic.

1.3 The History of Anatomy

Part one of this introduction will describe key points in the history of anatomy.

1.3.1 Ancient Egypt

The study of anatomy begins at least as early as 1600 BC. Most physicians of the time were also priests. Their medical practices were based upon religious beliefs, and it was common for physicians to specialise in a body area. This was usually because it was believed that different Gods governed different body parts. Medical practice was based upon a tradition of rituals. The oldest medical texts in the world date back to the Egyptians. Those still existing today not only describe human anatomy and disease, but also their treatments[8].

The Edwin Smith Papyrus is the earliest known medical document, written around 1700 BC. The ancient text lists 48 medical cases, and describes their examination, diagnosis and treatment [9, 10]. The papyrus contains the first descriptions of the cranial sutures, the meninges, the external surface of the brain and the cerebrospinal fluid. The document recognises the heart, vessels,

liver, spleen, kidneys, ureters and bladder and that the blood-vessels are connected to the heart. Some Egyptian practices, such as embalming, are still in use today.

1.3.2 Greek Medicine

The practices of ancient Greece (now Italy, Egypt, Greece and Turkey) are the foundations of medicine today. It was during these times that medicine moved away from the mystical and towards the logical and observed.

The first known physician was Hippocrates whose name and image began to emerge as a leader in medical research and thought during the 5th century B.C. Hippocrates was known for his reasoned thinking. He became best known for his use of hygiene, diet and drugs in healing. Key to his theories was the notion of keeping the body in balance. Illness was caused when fluids became out of balance, sometimes requiring the reduction in the body of a fluid through bloodletting or purging.

The Hippocratic Corpus was a collection of works attributed to Hippocrates. However, there are assumed to be several authors. Their writing is thought to have taken place over several centuries, hence the presence of often contradictory advice. The Corpus contains writing on a variety of medical

topics, including diagnosis, epidemics, obstetrics, paediatrics, nutrition and surgery, again fundamental to current medical practice[11, 12].

1.3.3 Galen

Galen (approximately 130 AD) was a medic in ancient Greece. As Roman law prohibited human dissection, Galen dissected animals including primates. Therefore a number of mistakes in recording human anatomy resulted as observations were from animal specimens[11, 12]. One famous example is the *rete mirabile*. The *rete mirabile* was described as a network of fine vessels, into which the carotid artery branched out at the base of the cranium[13].

Galen, renowned for his surgery, including eye and brain surgery, was noted to have performed cataract surgery. Galen, lectured, wrote extensively, and performed public demonstrations of his anatomical knowledge[12].

As a supporter of observation and reasoning, he was one of the first experimental physiologists. He performed vivisections of numerous animals to study the function of the kidneys and the spinal cord. His most common subject for dissection was the ape. Galen's work has informed much of the anatomy taught today.

1.3.4 The Opposition to Dissection throughout the Centuries

One medical school to achieve great eminence was the Alexandrian Medical School. It was the only school in the world where human dissection was performed regularly. The anatomists at the school were said to have performed human vivisection on condemned criminals. Herophilus was an Alexandrian medical scholar, said to have trained in the Hippocratic tradition. Herophilus dissected the human eye, skull, brain and liver[11, 12].

The rise in Christianity and its acceptance as the official faith of the Roman Empire during the fourth century reinforced the already common prohibition of human dissection. It was the belief of the church that the body was resurrected. In light of this belief, the church was firmly opposed to post-mortem examinations of any kind. Similarly, both Jewish rabbis and Muslims, also argued against the ‘desecration’ of the human body in later centuries[11, 12].

Due to the strong opposition to dissection it was Galen’s writings that formed the foundation for physiology and medicine in Christian Europe and the Muslim world until the sixteenth century[11, 12].

1.3.5 The Legalisation of Dissection

Emperor Fredrick II, founder of the Universities of Naples and Padua, fought with Papacy over the boundaries of monarchical authority. Early in the thirteenth century Fredrick ruled that all physicians in his domain were to learn anatomy by studying the human body. In addition, all physicians were required to produce documental evidence of such training.

These rulings applied only to Fredrick's holdings within Italy and Sicily. During the fourteenth century the papacy relaxed its prohibitations further. In 1341 the first documented post-mortem took place in Padua; 1348 saw the arrival of the Black Death[11]. The papacy then moved to approve post-mortems in order to establish the cause of the plague. However, it was not until 1537 that Pope Clement VII finally permitted teaching anatomy by dissection[12].

1.3.6 Early Anatomy Demonstrations

Early anatomy demonstrators followed Galen. Dissections were public events. Professors would read from Galen's text while a demonstrator pointed to the parts mentioned as a dissector did the cutting. The contradictions between the text and the anatomy revealed by the dissection were frequently ignored.

Likewise, Andreas Vesalius, one of the greatest anatomists of the sixteenth century and Professor at the University of Padua, was hesitant in challenging

what the authorities said on the basis of his observations. Consequently, the early works of Vesalius were repetitions of the Galenic tradition.

Artists of the time were responsible for many changes in thinking. Artists had no commitment to works of the established anatomical authorities and therefore drew as they observed. A prime example is Leonardo da Vinci's anatomical drawings, which have remarkable perception and accuracy [11]. Da Vinci produced around 750 drawings, but they had little impact upon medical progress. [12]

1.3.7 Vesalius and beyond

Andreas Vesalius (1514-64 AD) was a Flemish anatomist. Vesalius took his medical degree at Padua in 1537, and became Professor there. He later worked as a physician to the Holy Roman Emperor Charles V and to King Philip II of Spain[12]. In 1543 he published his masterpiece, '*De fabrica corporis humani' (Concerning the Composition of the Human Body)*. This text was highly illustrated. It confronted Galen's tradition and praised observation[11].

Vesalius' realisation that much of Galen's evidence had come from animal dissections rather than human helped to explain anomalies[11, 12]. The strength of Vesalius' work was that it came from direct observation and was well demonstrated. Although his work made no significant discoveries, it did

mark a significant change in method, moving towards a more observed and intellectual approach. Vesalius' work presented exact descriptions of the musculoskeletal and neurovascular system, as well as the viscera.

In 1561, Gabriel Fallopius published a volume of anatomical observations which both clarified and corrected Vesalius' own work. Fallopius was a student of Vesalius and his successor as professor at Padua[11]. Work by Fallopius included identification of structures in the ear and skull. More famously, he named the vagina, described the clitoris and discovered the tubes running from ovary to the uterus. These tubes later became known as Fallopian tubes, however Fallopius did not identify their function[12].

Ambroise Paré, a leading surgeon in the sixteenth century, used Vesalius' work as the reference for the anatomical section of his surgery text, published in 1564. Paré had translated much of Vesalius' work into French, thereby putting anatomy into the hands of the practising barber-surgeons.

Barber-surgeons were not only responsible for cutting men's hair, but also performed general surgery; often working for the British Army and Navy tending to the wounds of soldiers. The famous red and white poles of their barber's shops are still in use today. These symbolic poles stemmed from the imagery of merging blood and white bandages. It is reported that in 1745 the Guild of Surgeons became distinct from the Guild of Barbers in the UK. However, it was

not until 1800 that the Royal College of Surgeons received its Royal Charter[14].

By the end of the sixteenth century, Vesalian anatomy had entered the daily practice of surgery and medicine, becoming the standard for anatomical study across European institutions. This resulted in a new generation of anatomists, who began with an understanding of viscera, and the musculoskeletal and neurovascular systems. It was from here that the detailed work on the finer components of gross anatomy began[12].

The first detailed study of veins was published in 1603 by Hieronymus Fabricius. It gave accurate descriptions of the venous valves. Another student in Padua, William Harvey, published work on the circulation of blood in 1682. It provided an explanation of the purpose and function of valves. This was the first modern work of Physiology, and its anatomical content came directly from the Vesalian school[15].

Further discoveries continued throughout the seventeenth century. Gaspare Aselli of Padua described the vessels of the mesentery and identified their function of carrying chyle from food[12]. Aselli was the first to introduce colour to his work, this led to subsequent studies of the lymphatic system and the thoracic duct [11].

In 1644 the pancreatic duct was identified at Padua, as were the submaxillary duct (1656) and the parotid duct (1659). Towards the end of the seventeenth century anatomists had gained enough understanding of the accessory glandular system that it was possible for Franciscus Sylvius of Leyden to outline the chemical theory of digestion. Similarly, an accurate description of the reproductive system was produced by Regnier de Graaf.

1.3.8 Anatomy post seventeenth century

During the seventeenth century, microscopy was developed. This helped contribute to anatomical advances. It was Marcello Malpighi who took the lead in using microscopy for anatomical purposes.

Throughout the rest of the seventeenth and eighteenth centuries surgical training and teaching advanced significantly. Consequently, anatomy became more important. One of the reasons for this was the speed at which surgeons worked, due to the absence of anaesthetics, meaning a specialised knowledge of anatomy was pivotal. Surgeons such as John and William Hunter, Pierre Dionis and Antonio Scarpa continued to transform anatomy. The production of anatomical atlases began, often including memory aids such as triangles and quadrangles for identifying areas of surgical significance.

It was not until the late eighteenth century that demonstrated dissection became accepted practice. In the UK, formal demonstrations soon became accompanied by student dissection. In light of this the demand for cadavers increased. The demand was difficult to meet, especially with the laws restricting the disposition of bodies in England and Scotland. To keep up with this demand acquisition of cadavers often involved illegal activities such as body snatching and murder[15, 16].

The most famous incident of this involved William Burke and William Hare of Edinburgh. Burke and Hare provided bodies for dissections performed by Professor Robert Knox, all acquired by murder. The illegal acquisition of bodies by Burke and Hare was responsible for the Anatomy Act (1832), through which it became legal for bodies to be donated to medical schools if unclaimed, particularly in the case of those who died in prison or the workhouse[17]. Maintaining an adequate supply of cadavers still remained a problem for many medical schools. In her text '*Death, Dissection and the Destitute*'[16], Ruth Richardson suggests that the anatomy act permitting the use of unclaimed cadavers rendered dissection a punishment for poverty.

It was the mid-nineteenth century when gross anatomy became truly established. From then on work focused upon determining the fine structures[11], building upon knowledge acquired during the previous centuries.

The shift from large public displays of dissection, as was common in the 17th and 18th century, to classroom dissection led to a change in participation. For example, females were allowed to attend dissections. In the late 1850s there were a significant number of major scientific events. Charles Darwin published 'The Origin of Species' and the Medical Reform Act established the Medical Register and General Medical Council. Specifically relating to anatomy, the publication of the first edition of Gray's anatomy in 1858 was a major event. Wilhelm Röntgen pioneered the use of x-rays in 1895[15] (p314), an event which has transformed diagnostic medicine and anatomy teaching. Other significant developments followed over the decades, such as the creation of immunology in the 1900s, thanks to Louis Pasteur and Robert Koch. Technologies such as Computer Tomography (CT) and Magnetic Resonance Imaging were developed enabling medicine and anatomy to prosper.

The next section of this introduction describes anatomy teaching and current legislation.

1.4 Gross anatomy within the medical curriculum

"Clinicians often blame anatomists for teaching students too many details and not enough clinically relevant structures[18]."

In order to understand this accusation the anatomy curriculum needs to be considered in greater detail. Anatomy is taught in many ways. Traditional methods are didactic and include the students dissecting cadavers[19]. More

recently, medical schools have moved away from cadaveric dissection in favour of using prosections, living anatomy and plastic models. With technological advances and increased use of CTs and MRIs, it has become increasingly common for medical schools to integrate much more radiology and cross-sectional anatomy into their curricula.

It is not just the methods of teaching anatomy that need consideration. The structure of the entire undergraduate curriculum has seen dramatic change. Teaching has moved from traditional, teacher-led lectures and practicals, often with each subject (e.g. anatomy and pharmacology) taught as a unit in their own right, to more integrated approaches such as Problem-Based Learning (PBL) or case-led structures.

Since the publication of the General Medical Council's *Tomorrow's Doctors[20]* in February 2003, and more recently again in 2009[21], there has been a national change in the curricular structure. Two statements from Tomorrow's Doctors summarise the recent changes in the UK:

- "Modern educational theory and research must influence teaching and learning. Medical schools should take advantage of new technologies to deliver teaching[20]."
- "The clinical and basic sciences should be taught in an integrated way throughout the curriculum[20]."

Heylings (2002) [19] noted that in the period following the publication of Tomorrow's Doctors[20, 22] there has been a loss of gross anatomy teaching time. However, there has been greater integration between anatomical disciplines and clinical skills.

McLachlan *et al* (2004) stated that "anatomy learning is….seen as essential to medical practice…[23]." Despite this, evidence of the effectiveness of the various teaching approaches used is scarce. This has become evident from the literature review of articles relating to anatomy teaching.

1.5 Traditional methods of teaching anatomy

Traditional anatomy courses are often taught systemically. Dissection based practicals often proceed lectures. This is a didactic style of teaching.

Anatomy comes from the Greek word anatome, 'dissection', however, dissection based courses are now rare. Parker (2002) reported that dissection is now only compulsory in a minority of Australian medical schools[24], a trend mirrored in the UK.

Parker (2002) goes on to say, "anatomical dissection can be regarded as the impetus to the development of modern medicine[14]." High regard for dissection is shared by many authors and much debate persists about the advantages and disadvantages of dissection-based courses[4, 7, 24-34].

Although there is a considerable number of articles addressing the pros and cons of dissection, usually with the author advocating a preferred teaching method, the evidence supporting either argument is deficient. Yeager (1996)[34] called for more experimental evidence with regard to the educational benefits of dissection. Although this thesis does not address dissection, it does aim to provide evidence on the educational efficacy of other pedagogical approaches utilised in anatomy (chapters 3-5).

1.5.1 Advantages of cadaveric dissection

The proposed advantages have been divided into two broad categories, educational impact and personal and emotional development. The educational values category addresses the advantages which confer directly to developing the student's anatomical knowledge. Personal and emotional development addresses the advantages to the student which concern non-anatomical outcomes.

1.5.2 Educational impact

The educational impact of anatomical teaching using cadavers is widely debated. The extensive lists of educational advantages of using cadavers tend to stem from direct observations and a wealth of author experience. However,

most are assertions, lacking measures of comparison between various techniques.

The most common assertions published are:

Cadavers are excellent for the anatomy of large organs[24] and give an overview of spatial orientation[7, 24, 28]. They give students 3D view of human anatomy[23, 35, 36] and can reinforce and elaborate material acquired in lectures[35]. Body structure is reinforced in visual, auditory and tactile pathways when dissecting[7], and it confers the “feel” of tissues and organs[37]. Anatomical terminology is acquired in the DR, not by rote learning but by conceptualisation based on what is seen and felt.

Specimens are viewed as invaluable for practising surgical procedures[24] and use of surgical equipment[7] in a stress-free environment. They can be used for teaching clinical anatomy, e.g. intercostal drain insertion[24]. The study of human material is seen as an opportunity to appreciate the variability present in real human material [7, 35],and similarly gives an appreciation of whole-body pathology[24].

Cahill and Leonard (1997) quoted Dr. Marvin Wagner, “the computer is not a human[38].” This remark was offered in response to those who favoured teaching and learning anatomy by computer. Cahill and Leonard go on to paraphrase Wagner’s statement, “medical students share a moral obligation to

study anatomy by dissection; anything less is of questionable educational merit and may undermine the trusting relationship of future patients[38].”

In response to this statement, one might wonder whether or not patients question the anatomical training of a doctor, or if they are concerned about which means the doctor was taught? Do patients (assuming little or no medical knowledge) in fact realise that their doctor may have been taught by dissection or have studied from living anatomy, imaging, textbooks and plastic models? I would suggest that this is doubtful.

The paraphrasing by Cahill and Leonard continues, “anatomists must remain guardians of established educational principles, such as ‘learning by doing’ and ensure that anatomical instruction is facilitated, not compromised, by the computer[38].”

It is possible to deduce from this statement that perhaps Wagner is not concerned with the educational value of teaching by means other than dissection, but with preserving a tradition and anatomical ritual. This may mirror the often expressed anecdote that dissection is considered to be a rite of passage for medical students and that students who cannot stomach a dissection are not cut out to be a medic.

From a personal perspective, there is perhaps bias by anatomists for teaching in the way that they themselves were taught. Having studied anatomy by dissection as an undergraduate myself I am a strong advocate of this method. However, having moved to an institution where cadaveric dissection exists but is not the only method of anatomy teaching utilised I can also appreciate the merits of non-cadaveric teaching methods e.g. living anatomy and computer packages.

1.5.3 Personal and emotional development

It is not only the academic benefits of using cadavers that authors advocate, many authors express their experiences of non-academic advantages to students studying with cadavers. The cadavers are regarded by many as a medium by which to develop the student as a person i.e. as using the cadaver for personal development of the student in issues regarding emotion, professionalism and ethics.

Some articles suggest that dissection promotes humanistic values [1, 7, 23, 35, 36] and provides a platform for the teaching of moral and ethical issues[24]. Similarly, the cadaver has been seen to introduce students to death and trauma in a controlled manner [7, 35, 37] and provides a first acquaintance with the patient-physician relationship[7]. Working in the dissecting room (DR) has also been considered to be a good introduction to self-directed learning (SLD) and teamwork [7, 35].

1.5.4 Disadvantages of dissection

As one might expect there is minority of literature supporting non-dissection based teaching of anatomy. This literature states numerous asserted disadvantages of cadaveric dissection. The broad categories are as previously, however the costs to the institution are markedly high and are thus included by authors.

Before discussing the disadvantages of using cadavers, the difficulties in obtaining them warrants a mention. McLachlan and Patten (2006) highlight the fact that there has been, since 2001, a 10% reduction in the number of bodies accepted by Her Majesty's Inspector of Anatomy (HMIA) for anatomical studies[39]. The HMIA reported the number of bodies accepted in 2001 was 670, falling to 600 in 2004/5[39]. One reason for this decline was the health and safety directive forbidding the acceptance of donors diagnosed with mild dementia. This is an example of one of the ethical and legal difficulties that authors describe[23, 35, 36]. A more detailed discussion of such issues will follow later in the Chapter.

1.5.5 Hindrance to personal development

Understandably, one of the major hindrances in following a dissection based curriculum would be any negative reactions towards encountering death

experienced by students. A minority of students find the dissection process stressful and it may encourage inappropriate attitudes towards human remains as students attempt to deal with the stress[23]. Skidmore (1998) suggests that a fear of death may “cast an oppressive shadow and affect academic performance[40].”

1.5.6 Cost to institution

Multiple papers refer to the high cost involved in maintaining specimens and dissection facilities, as well as employing suitably trained staff. Articles by McLachlan [23, 35, 36, 39] are populated with observed disadvantages to the university for running a dissecting room.

A prominent thread, not only in the articles by McLachlan, is staffing. There has been a decrease in the number of teachers of anatomy, postgraduate students, clinicians and basic science lecturers[41, 42]. Also, suitably skilled staff must be appointed to preserve and dissect the specimens[23, 35, 36]. Even if the skilled staff are recruited Plack (2000) [43] remarks that the specimen produced is only as good as the skill of the dissector. A further limitation suggested by Plack (2000) is that any mistakes made whilst producing a prosection cannot be undone[43].

Both Plack (2000)[43] and McLachlan (2004, 2004, 2004, 2006) [23, 35, 36, 39] proceed to note the costs involved for an institution for simply having a dissection facility. The potential costs associated with the appropriate storage of the specimens in accordance with the UK Anatomy Act (1832,1984) and the Human Tissue Act (2004)[44, 45] are considerable. The problem of storage is two-fold, firstly the previously mentioned costs, but secondly and more importantly is the decrease in quality of the specimen once in storage (Plack, 2000)[43]. Storage and dissection facilities must be adequately ventilated. Flow

rates for ventilation in DR must be very high; their achievement is very expensive[35].

Similarly, there are the large costs of transporting, maintaining and disposing of cadavers[39]. Preservation of the cadaveric material is expensive. Of paramount importance, with respect to preservation of the specimens, is the potential cost to the health of the students and staff working with them. Embalming may fail to eliminate potential health risks such as AIDS and prion diseases in the cadaveric material[35]. Berube (1999) raised concerns over health risks involved in spending extended periods of time over formalinised wet specimens[46].

Once all of the issues regarding obtaining and storing the bodies, staffing and maintaining facilities are overcome, the institution then needs to consider the educational value of the cadaver. As we have seen, many authors are strong advocates of learning anatomy through dissection, however, as expected, authors are aware of the limitations of the method.

1.5.7 Lessened educational impact

After the embalming process, fixation alters the colour and texture of tissues[35], the texture, colour and smell of the cadaver are no longer real to life[24, 35].

Plack (2000) points out the specimens only “portray the body in a limited and static manner[43]”, a comment supported by McLachlan (2004) when he describes how cadavers are not a good guide to living anatomy as they are not responsive to movement and cannot be asked to change position, nor can they participate in interactive investigations (e.g. auscultation, palpation or percussion[23, 35, 36]).

With respect to anatomical structures and pathology; cadavers tend to come from a similar population: the elderly suffering from degenerative changes. Students relying on cadavers may therefore be basing their knowledge on something that does not represent ‘normality’[24]. Nor, are they ideal for studying skeletal anatomy ,nervous system anatomy (especially small nerves), surface anatomy, anatomy of small or indistinct organs (e.g. parathyroids, pineal gland, pancreas), anatomy of lymphatics and muscular anatomy in the contracted state[24]. Furthermore, the information gathered from dissection does not readily translate into cross-sectional views[35].

Although all of the points raised are valid, many can also be said of the suggested materials for replacement of cadavers e.g. plastic models. Such models are extremely expensive, prices ranging from hundreds to many thousands of pounds. Plastic models do not have a real to life texture or colour. Often the models are colour-coded to denote various structures, making it

obvious for students which structure is a nerve. Thus, they are not always ideal for use in practical assessment.

1.5.8 Dissection versus prosection

The hierarchy between dissection and prosection, specifically with regard to educational impact, is much debated [34, 40, 47-49]. I have previously presented the perceived advantages and disadvantages of following a dissection based course.

When considering dissection versus prosection, two schools of thought exist; it has been argued that the hands-on approach of dissecting enhances learning and provides students with confidence in the subject matter[49], and conversely it is suggested that using prosections provides students with the benefit of learning from cadaveric materials but without the emotional distress the dissection process can cause[40]. The latter is further supported by Dinsmore[50] who described the efficacy of using prosections.

Due to the contraction of hours provided for gross anatomy teaching, many institutions have moved away from the classic, full body dissection to varying degrees of prosection. It is reported this reduces the amount of contact hours while allowing students the exposure to structures that they might otherwise spend hours trying to find[50]. Alexander (1970) found no significant differences in outcomes between students learning anatomy from dissection relative to those learning it from examining prosections[51]. Nnodim and Yeager support an increased use of prosection[34, 47]. They suggest prosections as a means

of maximising utilisation of scarce materials while increasing educational efficiency, as measured by outcomes on examinations and long-term recall.

A solution to limited time for dissection has been the introduction of peer-led teaching[52, 53]. Students alternate with their peers in taking responsibility for dissection of their assigned cadaver. The designated dissectors then teach their peers about the region. Nnodim (1997) evaluated peer-led teaching, finding that it was well-received by the students. Students reported that it reduced the crowding at dissecting tables considerably and offered them opportunities to hone their communication skills. Additionally, the non-dissecting time was also valuable and utilised by students[52].

One can imagine that there might never be a conclusion to the debate surrounding dissection and prosection, especially given that anatomy is a discipline steeped in tradition. However, due to curricula restrictions, the use of prosections remains popular and I believe that this will persist.

When considering the potential advantages and disadvantages of working with human material, one cannot afford to overlook the Human Tissue Act.

1.6 The Effect of the Anatomy Act (1984) and the Human Tissue Act (2004)

The Human Tissue Act (2004)[44, 45] arose after the *Kennedy[54]* and *Redfern[55] Inquiries* into improper storage of organs and tissues from children who had died at the Bristol Royal Infirmary and the Royal Liverpool Children's Hospital (Alder Hey) during the period of 1999-2000. These inquiries found that the organs and tissues had been removed, stored and used without proper parental consent. The *Isaacs Report* (2003)[56] found that this was a widespread practice and that the law on this area was neither clear nor comprehensive.

The Act now covers issues regarding whole body donation, and the taking, storage and use of human organs and tissues. It regulates activities such as post mortem examinations and the storage of human material for education, training and research.

"It is intended to achieve a balance between the rights and expectations of individual and families, and broader considerations, such as the importance of research, education, training, pathology and public health surveillance to the population as a whole[44]."

The Anatomy Act (1984) stipulated that anyone intending to practice anatomy must obtain a licence from the Home Secretary. Teachers who take out such a

licence accept responsibility for the proper treatment of all bodies dissected in the building for which the licence was granted. Inspectors were employed to regulate these licensed teachers and ensure that all cadaveric material was accounted for.

With such strict regulations, I would suggest that Institutions are moving away from dissection based courses for ease, cost efficiency and due to time restraints. It is easier to store a plastic model than a wet specimen and no licence is needed.

A valid point raised by Forsyth *et al* (2006) in an article about dentistry refers to the nature of the material covered by the Human Tissue Act[57]. As outlined above, the act covers the removal, storage and use of 'relevant material' from the deceased, as well as the use and storage of 'relevant material' from the living. "The 'relevant material' is defined as all material that consists of, or includes, human cells other than gametes, embryos outside the body, and hair and nails from a living person. The Act therefore covers human teeth, both from the living and deceased[57]." It is when you consider that even teeth are covered by this Act that you realise the magnitude of its potential impact on medical education. Licensing, inspections and added cost are all feasible deterrents for medical schools not to use cadaveric material.

A further impact of the Act is that of its effect on the number of bodies available for use. The Act implemented a new legislation that formal consent to bequeath a body not only had to be written, but also witnessed. This was one of the prompts for Sir Liam Donaldson (Chief Medical Officer) to write in 2006 to all doctors in England warning of a shortage of bodies and asking them to encourage patients to bequeath their bodies for medical training and research [58, 59].

After the negative press of the *Kennedy* and *Redfern Inquiries*, there was a great deal of public mistrust in the National Health Service and associated organisations. Such bad press will undoubtedly have contributed to the fall in body donations.

1.7 Further Public Mistrust?

Channel 4 television screened a series '*Anatomy for Beginners'* in which Professor Gunther Von Hagens and Professor John A. Lee performed dissections in front of a public audience. The same network aired a second programme '*Autopsy: Emergency Room'* during November 2007 due to the high public interest in *Anatomy for Beginners*.

One may wonder what was it that got the public so interested? There are many factors which may have contributed such as the taboo surrounding the subject, the gore factor or a simple interest in human form and function. Whatever the

ingredient that raised high public interest, the show was much talked about, both by the public (often due to the nudity of the models) and by the scientific and medical world. Regarding the latter, the production was not as well received as the producers may have hoped.

In a personal communication, an anatomist colleague suggests the show as having a negative effect on the number of body donations in the UK, due to the bad press it created[60].

1.8 BODY WORLDS™

The BODY WORLDS™ exhibitions[61] are a collection of real human specimens which range from whole body plastinates to transparent body slices. Exhibits aim to demonstrate the effects of bad health and lifestyle choices to the visitors. The professed aim of the exhibition is to teach the public about anatomy, physiology and health. *“The primary goal of BODY WORLDS is health education[61].”*

The plastinates within the exhibition have been posed to demonstrate muscle function, for example, when playing sport. The idea of this is to allow the public to relate the plastinate to their own bodies.

The first exhibition was held in Tokyo, Japan between 1995 and 1999. There were 2,945,825 visitors. To date there have been over 25 million visitors. Professor Lantermann from the University of Kassel, Germany, developed and conducted a survey polling visitors reactions to BODY WORLDS™ exhibitions. The survey was conducted across many cities in different continents; therefore it is likely to have included multiple cultures and beliefs[61].

Lantermann's survey[61] concluded that only 5% of visitors felt that the display of human specimens offended their views on human dignity but 66% of the visitors questioned reported leaving the exhibition with 'valuable incentives for a healthier lifestyle'.

So why are the opinions of visitors to a museum exhibit important to the teaching of anatomy in a Medical School?

Firstly, it is the public's cadavers we are relying on for donation in order to continue with dissection based programmes. If exhibitions such as BODY WORLDS™ get good feedback from the press, public and medical colleagues then the public will look more favourably on dissection, especially in light of the controversy of the *Kennedy* and *Redfern Inquiries[54, 55]* in the UK.

Public observation of the cadavers being treated with respect and being well maintained will evoke more positive responses to the work of medical educators, researchers and anatomists, in turn, promoting the idea of body donation.

In Lantermann's survey, 20% of the visitors stated that they 'could imagine donating their body for plastination purposes after death'. This statement may not show any level of support for body donations to medical schools, however. This may be because people wishing to have their body plastinated may want it to be immortalised, or may be utilising plastination as a means to avoid funeral expenses.

Secondly, public support is needed for many reasons other than body donation. For example, financial donations to support research and willingness to have opinions and medical information canvassed.

Exhibitions such as BODY WORLDS™ bring anatomy, formerly kept behind closed doors, to the public arena. This shows the public that practices are legitimate. It may also elicit an interest in anatomy within the public. Those interested may then wish to help in Medical Schools as life models and clinical skills partners. This may, of course, not only be to assist the students but to aid their own education, even just as a hobby. Conversely, with televised dissection

and plastinated cadavers on display in museums one might anticipate a public backlash.

In Winter 2005, a Scottish T.V. network interviewed Professor Sue Black and Dr. David Dawson from the University of Dundee's Anatomy and Forensic Anthroplogy Department about their views on the show and the BODY WORLDS™ exhibitions produced by Gunther Von Hagens.(BODY WORLDS is a travelling exhibition of plastinated specimens dissected at Von Hagen's Institue of Plastination) Professor Black remarked that while Gunther Von Hagens was clearly very talented, his work had an element of showmanship, with which came disrespect for the donors on whom he was demonstrating.

Internet search engines retrieve tens of thousands of articles referring to BODY WORLDS™ exhibitions, many in favour, equally as many against. A common theme in articles criticising the exhibition is a feeling that the human dissection has been 'sensationalised and trivialised[61].'

Literature suggests that "the public display of cadavers, by providing an adroit interface between the craving for sensation and acquisition of comprehensive information, could actually contribute to improved health-consciousness and awareness of one's own mortality[62] (p572)." Exhibitions such as BODY WORLDS™ have been reported to have a positive effect on learning, with

visitors reporting feeling more informed about their own body. It is the vividness and authenticity of the specimens which have a positive impact on visitors[63].

Advocates of exhibitions such as BODY WORLDS™ are for public display of cadavers for a variety of reasons. Firstly, plastination represents the modern history of anatomy and should therefore be reported. Plastinates are no longer viewed as bodies but as models and therefore should be displayed. Visiting the exhibition is voluntary so people are making a conscious decision to view the displays. And finally, the educational benefits are advocated[62, 63].

Contrary to this are the views that exhibitions of this nature should be prohibited. Concerns have been expressed over the dignity of the cadavers. This has been related to the poses, such as cadavers playing poker, and to the sexualisation of the body due to an emphasis on the genitalia and sexual organs. The gestalt positions of the cadavers has also been criticised for sensationalising the cadavers. However, it is this positioning which is reported to generate much of the public interest. Finally, the permanent anonymity of the cadavers has been criticised as it results on complete medicalisation of the body and removes all traces of identity and association[62, 63].

Research has suggested that the disproval of BODY WORLDS™ may be associated with burial culture[63]. This is because laws relating to burial have

evolved from tradition and social assumptions. Exhibition of cadavers is perhaps challenging these social expectations.

Researchers have found that medical and nursing students attend such exhibitions as they feel that it is expected of them, as well as visiting for educational benefits[63].

1.9 Other methods of teaching anatomy

Dissection is likely to be the most well known and historic method of teaching anatomy. Records from around 140-180 AD note animal dissections being performed by Galen[11, 12]. Although dissection is still popular, technological advances have allowed teaching to move towards electronically based media. Examples include the Visible Human Project[64] and the resultant software such as the Virtual Human Dissector™ , the latter of which will be discussed in later Chapters.

1.9.1 Electronic media

Many electronic anatomical resources exist. These include; DVDs such as Acland's Anatomy©[65], which presents cadaveric prosections with audio commentary and websites such as Instant Anatomy™[66] which provide diagrammatic summaries of regional anatomy. Such media are advantageous to students as they allow access to materials outside university hours. These

electronic resources, particularly those which use diagrams or computer generated imagery, provide a solution for students who may have emotional distress when confronted with cadavers.

Electronic media provide three dimensional (3D) representations of anatomical structures which can be manipulated by the user to provide views which may not be possible within the laboratory. Images are of high resolution and quality. Images can be reused, reuse is limited with cadavers and prosections due to a decline in quality over time. Electronic media also support distance learning. These issues are described further in Chapter 3.

1.9.1.1 The Visible Human Project™

The Visible Human Project™ is a complete, anatomically detailed, three-dimensional representation of an entire male and partial female human body. A male and a female cadaver were both cut into thin cryosections which were then photographed and digitised. CT and MRI scans were also taken of both cadavers. The male donor was a 38-year-old Texan murderer who was executed by lethal injection. The cadaver was frozen in a gelatin and water mixture in order to facilitate sectioning. The specimen was then cut in the axial plane at 1 millimetre intervals. Both analog and digital photographs were taken of the resulting 1871 slices[64, 67].

The 59 year old female donor remains anonymous. The female cadaver was cut into 5189 slices at 0.33 millimetre intervals. She was thought to be a Maryland housewife, who died of a myocardial infarction.

The resulting digital image library allows for identification and multilingual labelling of every visible structure in the data set[64, 68]. Researchers are now starting to describe the efficacy of such data sets in anatomical education[67, 69].

Many software packages have evolved from the Visible Human Project DataSets™. Examples include: NCI Flicker, which compares images from different internet sources, and the Segmented Inner Organs Data of the Visible Human Male.The Virtual Human Dissector (VHD) is another interactive anatomy programme based upon that very dataset. [64, 67]

The educational value of the Virtual Human Dissector will be discussed in Chapter 3. At this point in the thesis I am acknowledging that such databases as the Visible Human Project™ and related software such as the Virtual Human Dissector™ exist and serve to make access to cross-sectional anatomy more readily available.

1.9.2 Living Anatomy

Living anatomy is the observation of the living body as opposed to the observation of cadaveric anatomy. Living anatomy, also described as surface anatomy (although a distinction exists), is the anatomical form in which most medics encounter the body. In practice, medics encounter the patient as a living and moving being. This is in stark contrast to the cadaver, from which most medical students learn their anatomy. Cadavers, as discussed earlier, are non-motile and unresponsive, and differ in colour and texture [23, 43].

Literature recording the teaching of living anatomy dates back to books by Robinson (1928)[70] and Lockhart[71] (1947) describe and depict living anatomy. Some of the earliest recordings of living anatomy date back to the eighteenth century and John Hunter[15].

Recommendations for the teaching of living anatomy were made by the General Medical Council in 1930, “the demonstration of structure and function in the living” should form an integral part of medical students’ professional training (cited in Waterson 1931)[72].

Living anatomy teaching sessions can take many forms. Current methods include peer-physical examination, body painting and various types of imaging and scanning. Techniques such as imaging and examination are utilised within

the practitioner's setting; examining a patient is an exercise in living anatomy. The focus within this introduction is on living anatomy as a teaching method for undergraduates.

One of the biggest advantages of this teaching methodology is that anyone can participate. As the cadavers permitted for use must now be only those having died from natural causes (Human Tissue Act 2004)[45], the pool of donors has declined and the bodies tend to come from older donors. Students are not seeing an accurate representation of the population[24].

For living anatomy sessions anyone able to consent can participate as a life model, sometimes referred to as a clinical skills partner[73]. This means that the students are able to see a variety of body morphologies and a wider variety of medical conditions. If the students participate themselves, by examining their peers they get to see generally younger and muscular physiques than those of the cadaveric population. Another obvious advantage is the ability of those being examined being responsive to commands and touch, whether it be in a clinical examination or a body painting scenario.

1.9.2.1 Surface Anatomy

Surface anatomy is the part of living anatomy specifically concerned with the structures that can be identified from the outside of the body, often, for

example, bony landmarks.

Surface anatomy is the form in which most doctor's first encounter their patient's anatomy. For example, if one was to listen to the lungs of a patient they would need to know where to place the stethoscope on the chest. To work this out the doctor uses surface anatomy. It is in clinical examination that students apply the anatomy taught to their patients, specifically through living and surface anatomy.

Barrows *et al* (1968)[74] remarked that medical students were often not given the chance to transfer the anatomy that they had learnt within the laboratory directly onto a living body. The authors call for students to be able to palpate structures and see the structures studied in the cadaver move and function; "In this manner a solid learning bridge is formed between anatomical knowledge and his further pursuits in physical diagnosis and clinical medicine is formed[74]."

Barrows *et al* (1968)[74] draw parallels between the reactions of the students towards cadaveric dissection and examination of a patient. The students generally have an objective yet respectful attitude towards the cadaver, this objectivity is not always mirrored when faced with a disrobed human body. The authors note that embarrassment not only impedes the students' clinical learning but can also affect their subsequent rapport with patients, and any

unspoken uneasiness shown by the student when examining can adversely affect the patients' responses.

The article[74] makes reference to peer-physical examination as previously being the only attempt to correlate living anatomy with the rest of the anatomical syllabus. The peer examination was viewed as being 'unsatisfactory' for a variety of reasons. For example, embarrassment as disrobing among peers was viewed as inhibitory, especially if examiners and examinees were of opposing sex. Embarrassment, however mild, both inhibits observations and 'precludes adequate palpation and probing by examiners'.

Barrows[74] proceeds to discuss the role of the doctor-patient relationship. The authors view this relationship as one of submission by a patient in order to facilitate an intensive examination and thorough pursuit of all facts necessary for diagnosis by the doctor. The authors suggest that students ought to start their experience of living anatomy with a similar professional relationship, and thus not examine fellow medical students[74].

The patient is also seen as an unsuitable subject for students to practice upon. Examination by the students could be both tiring and uncomfortable. Barrows *et al* (1968)[74] suggest that paid models are a suitable alternative as they would expect exposure and repeated handling by the students. Similarly, the student would not fear discomforting someone who is not ill. The above was echoed by

Stillman *et al* (1978)[75], who added that repeated contact with a partially robed or disrobed model may hasten then student's comfort in actual clinical situations with patients.

Both Barrows *et al[74]* and Stillman *et al[75]* conducted student satisfaction surveys after running 'clinical correlation' anatomy sessions where students examined models under the supervision of clinicians. Both studies found that students highly rated the models and the surface anatomy taught. For example, Stillman[75] found that 93% of students rated the sessions as being either 'very good' or 'good'. Perhaps the most memorable finding from Stillman *et al[75]* was that students rated the life models as superior to using cadavers, especially for purposes of superficial anatomy and landmarks.

Surface anatomy is a way of bringing cadaveric dissection to life[76]. A study by Aggarwal *et al[76]* found that when students were requested to be the life models, most were reluctant and only did so if no one else did first. The majority of volunteers were male. Students stated reasons ranging from embarrassment to finding being the volunteer a disadvantage when it comes to learning the material covered. Same sex tutorials did not increase willingness significantly, however this did reduce the degree of embarrassment experienced by some students. These problems can be addressed using life models; however, it is both costly and difficult to organise, difficulties also acknowledged by Stillman *et al* (1978)[75]. Aggarwal *et al[76]* make a good point that when there is a reluctance for students to volunteer the interactive

session becomes nothing more than an illustrated lecture. This study once again echoed previous findings by Stillman *et al[75]* and Barrows *et al[74]*, that student disrobing in front of peers leads to embarrassment and a lack of professionalism, as well as inhibiting the performance of adequate examinations.

1.9.2.2 Peer Physical Examination

As illustrated previously, how anatomy ought to be taught is a subject in hot debate, with authors predominantly supporting dissection or the use of prosections. Alternatively, there are rare cases where the author is a strong advocate of an alternative, such as McLachlan *et al* (2004) who concluded that anatomy ought to be taught using medical imaging and living anatomy [35].

Alternative methods for incorporating living anatomy include the use of life models or peer physical examination could be used as alternatives.

Peer physical examination (PPE) uses student volunteers for the practice of clinical examinations [77]. PPE bears many similarities to the use of life models; namely that students can assuage their own anxiety in performing examinations through repeated practice in a safe environment and they can view a wide range of body morphologies [77] (although perhaps not as wide ranging as may be possible through use of life models as the majority of the students will be

young and healthy). One advantage of PPE is that it is less expensive than using life models and easier to organise. Also, by being examined themselves, students appreciate what it feels like to be examined [77]. This will assist them in understanding the anxieties a patient may experience.

Rees *et al* (2004)[78] explored the attitudes of first year medical students towards PPE. They found that 97% were comfortable with participating with PPE of all body parts except breast and inguinal regions. Over 20% were unwilling to participate in PPE of the breast and inguinal regions. The students were more comfortable with PPE within gender rather than across. Females were the most likely to be uncomfortable with PPE. There were no statistically significant differences between attitudes towards PPE and age or religious faith. The authors argued that this might be a type II error and recommended further research.

From my own experience of assisting in PPE classes there is a correlation between religious faith and willingness to participate in PPE.

In a subsequent paper, Rees *et al* (2005)[79, 80] raised the point that the majority of literature surrounding whether demographics influence willingness to participate in PPE is of a quantitative nature, and does not explain the reasons behind such attitudes and relationships. McLachlan and colleagues have explored changes in student willingness to participate in PPE and have made

subsequent recommendations for optimising participation, such as allowing self-selection of peers with whom students will conduct examinations[81].

Rees *et al* (2005)[80] found that students rated PPE as valuable for many reasons. PPE developed their anatomical knowledge. In particular, the ‘hands on’ approach facilitated their learning of surface anatomy. The students commented that PPE gave them an understanding of variation in people, and what ‘normal’ feels like. It also aided the students in carrying out examinations without feeling embarrassed and provided a fun environment where students could make mistakes and improve their knowledge and understanding. They reported that PPE allowed them to empathise with their patients with regard to the fear and embarrassment of being examined.

The most common concerns surrounding PPE were that students felt awkward examining peers, specifically regarding the breast and groin. There was concern over not knowing if a peer had any sensitive issues that the examiner was unaware of, and that they may physically hurt their peer when practising.

Concerns about being examined were mostly related to intimate examinations. The responses collected by Rees *et al* (2005) ranged from students feeling embarrassed to ‘invaded’ or ‘vulnerable’, the reasons cited for this were that the area was ‘ too personal’ or feeling self-conscious. Concerns over self-image were negative and included anxieties about weight, scarring, lumps and bumps

and other physical imperfections. Some students also had visible medical conditions, one student in particular expressed a concern over this, “I have a [visible medical condition]. Although I don’t have a problem with it, I feel that other students may[80].”

As mentioned previously, Rees *et al* (2005)[80] did not find any relationship between willingness to participate in PPE and religious faith. There were comments made regarding this, “(PPE) is a good teaching tool but I don’t personally want to be examined because of religious reasons” and a further comment from another student, “because of my religion I don’t want to expose my body unless extremely necessary.[80]”

From the existing literature it is possible to deduce that, for the most part, it is females who are least likely to be willing to participate in PPE. Students are keen to engage in PPE, although some are unwilling to participate in PPE of the breast and inguinal regions.

The challenge for anatomists is therefore to find way for all students regardless of gender, religion or morphology to participate in PPE, and feel comfortable whilst doing so. Qualitative analysis of student satisfaction surveys doesn’t shed much light on the surrounding issues.

As previously described McLachlan *et al* [81] have researched the underlying issues that students' have with regard to PPE in their paper "Student attitudes to peer physical examination: A qualitative study of changes in expressed willingness to participate."

Prior to McLachlan *et al* (2010)[81] research has proposed that a student's willingness to participate in PPE is not static. Recent evidence[82] suggests that students' willingness to engage in PPE was stable before and after exposure to this teaching methodology. The observed changes included a slight decline in the willingness to examine, a slight increase in the willingness to be examined, with females being more likely than males to report a change in willingness (in both directions).

The approach of McLachlan *et al* (2010)[81] was to conduct focus groups to explore the reasons behind the students' changes in views, focusing primarily but not solely on female students as once again they were found to be most likely to change their mind.

Themes emerging from the McLachlan *et al* (2010)[81] *Grounded Theory* approach mirrored the free text comments collected by Rees *et al* (2005)[80]. An interesting new theme that emerged was that of the impact of staff upon a student's willingness to participate. One student reported that staff expectations and prejudices proved inhibiting in her case, "*a lot of the tutors were like 'ah,*

don't worry girls, we won't ask', I think their response influenced me by thinking, I was thinking 'oh it must be bad'. Do you know what I mean? It must be embarrassing if they're having, if the tutors are having to reassure us constantly as a year group then there must be something wrong with it …. and I thought, 'ah well, there must be a problem then.' I don't know, I think it's, tutors' attitudes can influence the way we feel as well[81]." From this it was concluded that, "it was staff barriers rather than student barriers that were significant in inhibiting participation in peer physical examination."

The authors found student preference for group composition to be dichotomous, "some students were plainly more comfortable with friendship groups- others preferred strangers[81]." This apparent division within the cohort indicates that there may not be a single successful strategy which promotes willingness to participate in PPE, as highlighted by the authors.

Furthermore, it was reported that the majority of students who changed their mind (in either direction) regarding willingness to participate in PPE were female. This mirrors previous research by Rees[77, 78, 80, 82]. In 2007, Rees[77] investigated the practical implications of feminist theory of body image and its relationship with PPE.

Rees (2007)[77] notes that the most consistent finding across the existing PPE literature is that "students' attitudes towards PPE are associated with their

gender. Male students are more comfortable with PPR than females and students are more willing to engage in same- rather than mixed-gender PPE."

In order to explain this relationship Rees[77] describes the feminist theory of Objectified Body Consciousness (OBC). Objectified Body Consciousness theory describes how females come to view their bodies as observable objects through 3 mechanisms: body surveillance (viewing one's body as an outside observer), the internalisation of cultural body standards leading to body shame, and appearance control beliefs.

The definition of body images used by Rees is that by Grogan, "a person's perception, thoughts and feelings about his or her body[77]." Rees cites Spitzack and McKinley, who explain that "girls learn quickly that they are appraised by others on the basis of their appearance and, consequently, come to experience their bodies in term of how they look to others[77]." In contrast to this male bodies are often viewed differently, in terms of functionality as opposed to aesthetics.

It is now worth considering how Objectified Body Consciousness theory relates to PPE. During a PPE session the students closely scrutinise the bodies of their peers, as pointed out by Rees[77], this is visual, tactile, auditory and olfactory. Examination often requires exposure of body parts through undressing.

We can recall from the theory of Objectified Body Consciousness that females have higher levels of body shame and surveillance compared with men. During a PPE session Rees[77] suggests that this may be exacerbated as the students are forced to watch their own bodies and the reactions of others toward their body. Rees[77] describes Objectified Body Consciousness as a social and dynamic process rather than an individualistic and static state. Rees also cites Tantleff-Dunn and Gokee, individuals' interpersonal relationships influence how they feel about their appearance through various mechanisms, including feedback on their appearance and social comparison.

The conclusions drawn by Rees were that females may be proportionately less comfortable with PPE than males because they have higher levels of body shame and body surveillance. They may also be less comfortable with mixed-gender PPE because they perceive the males to be the perpetrators of critical and teasing comments and feel at risk of sexual objectification by males[77].

My emphasis on PPE within this introduction comes from my firm belief that clinical skills and anatomy teaching must be integrated within the medical curriculum. This topic is further discussed in subsequent chapters. However, peer-physical examination allows for their integration at an earlier stage of undergraduate training, when access to patients might not be possible. PPE is fundamental to many other innovations in anatomy teaching, such as body painting and the incisions gown, therefore an understanding of PPE and

appreciation of its acceptability with students is paramount to its success and to our understanding of anatomy teaching within this introduction.

1.9.3 Art and humanities in anatomy teaching

In January 2008, Professor John McLachlan from Durham University unveiled the ‘incisions’ gown[83]. The gown was a collaborative project with The University of Ulster and Durham University, funded by the Wellcome Trust[84], produced as part of a wider project exploring the use of art in teaching.

The gown, now used in teaching at Durham, has a series of zips demonstrating incision sites for various surgical procedures. The gown can be worn by the students during demonstrations. It serves multiple purposes: to demonstrate the incision sites to students, to evoke an emotional response to incision and potential scaring from the students and it serves to bring anatomy and art together[83, 85].

Under the outer layer containing the zips are multiple layers of a rubberised fabric. This fabric was chosen as it resembles the texture of muscle fibres. The fabric has been arranged to mimic the muscle layers found under each zip, thus indicating to the fibres that must be parted in order to perform the procedure in question.

Within days of the official press release from Durham University the gown had received much acclamation in the press. It featured on two television networks' local news stations and was the subject of many newspaper articles.

An article from the BBC webpage quoted Professor John McLachlan, creator of the gown, as saying, "Current anatomical teaching aids describe but they don't evoke. They take no account of emotional involvement or the feel of the body. The way medical students distance themselves emotionally from the patient's body has long been seen as a desirable outcome of current modes of medical teaching. But this 'desensitisation; also brings with it the risk of objectifying the body. The patient becomes 'the liver in bed four' rather than Mrs Smith'[83]."

The purpose of describing this gown within my introduction is to highlight how some institutions are in their approaches to anatomy education and how they are responding to the decline in body donors. Innovations such as the incisions gown emphasise the clinical relevance of anatomy.

1.9.3.1 Body Painting

Body painting is considered by some to be the most ancient form of art. The skin is painted temporarily, but in some cases the designs can last up to weeks. Origins of body painting stem from tribal cultures and are ceremonial[86, 87].

Body painting is now most frequently seen on the face, usually at sporting events or children's parties. Although, many artists use the body as a canvas. Body painting is now being used in the medical setting as a way of projecting the inside on to the outside.

Body painting is an innovative method of anatomy teaching which encompasses peer-physical examination and surface anatomy, which have previously been described.

Op Den Akker *et al[88]* first used body painting in teaching in 1999. This session was 3 hours long. The students both painted and were painted, and also practised physical examination skills appropriate to the session (either abdomen or thorax) such as auscultation and palpation. Senior students were trained prior to the session and acted as supervisors. All participants worked from a manual produced by staff in collaboration with a professional body painter[88].

The results of Op Den Akker *et al [88]*are worth noting. 240 students participated in body painting sessions, but not one student complained about any technical difficulties. 67% of the students who responded to the survey issue did not report any feeling of embarrassment due to partial nudity.

The paper reports that using this method with large classes is both feasible and motivating. Body painting was efficient in terms of time and learning. It also appeared to be relatively inexpensive, as commercially available paints were used.

Op Den Akker (2002)[88] described body painting as painting internal structures on the surface of the body with high verisimilitude. McLachlan and Regan De Bere (2004) [36] and Sugand *et al* [19] support the efficacy of using body painting in conjunction with clinical skills, as described by Op Den Akker (2002).

Figure 1: Second year Medical students at Durham University showing their face painting from an anatomy teaching session.

1) Sensory innervation of the Trigeminal Nerve and the path of the Facial Nerve
2) Sensory innervation of the Trigeminal Nerve and the path of the Facial Nerve
3) The parotid gland, facial nerve, obicularis oris and obicularis oculi.
4) Sensory innervation of the Trigeminal Nerve.

It is claimed that the value of body painting as a teaching tool was partly due to its kinaesthetic nature[87, 89, 90]. In order to heighten the visual impact McMenamin suggests that the colour of the paints ought to correlate to the structure being painted as accurate as possible[89].

1.10 Conclusions: anatomy teaching

This introduction to anatomy teaching has highlighted how approaches to anatomical education have been shaped by the past, with practices of embalming and dissection dating back to the Egyptians and Ancient Greeks.

In response to the current climate within the UK education system, as well as technological advances and student expectations, innovative approaches to anatomy teaching have been pioneered. The efficacy of some of these approaches, such as use of the Virtual Human Dissector and body painting will be explored in Chapters 3 and 4.

1.11 Anatomy and Professionalism

Previously this introduction has focused upon anatomy, with particular emphasis on its teaching and history. However, medical students not only need to become competent physicians, with a strong scientific foundation, but also need to practice using these skills in a professional manner.

At this point it would be fair to ponder how anatomy and professionalism are related. Gregory and Cole (2002) best demonstrate this association with the following statement, “anatomical dissection in…medical education combines the ideals of the profession…. acquisition of scientific knowledge and skill balanced by the development of humanistic attitudes and behaviours.”[91] It is these humanistic attitudes and behaviours that define the professional behaviour expected of medics.

Before discussion of the attributes that a professional physician is expected to possess it is worth noting that Gregory and Cole [91] are not alone in drawing parallels between the dissecting room environment and professional values.

Granger (2004) notes that the anatomy course educates students in psychosocial areas critical to the development of the whole physician[7]. I have previously referred to articles which propose many advantages of a student following a traditional course in cadaveric dissection. Such articles suggested that these courses introduce students to the patient-physician relationship and concepts of humanistic care [7] and provide opportunities for teaching ethical and moral issues [24].

Having described how the teaching of the anatomical content of the medical curriculum can emphasise the principles of professional behaviour in medical students, what exactly constitutes professionalism should be considered.

1.12 Professionalism

This introduction to professionalism will discuss the constituent elements of the concept and how professionalism is assessed in an undergraduate setting. The purpose of this introduction is to provide an overview of the assessment of professionalism as experimental chapters 6 to 9 are related to this topic.

1.12.1 A definition of professionalism

When reading the literature on professionalism in medicine it becomes evident very quickly that there is not one clear and concise definition of professionalism [92-94]. There are extensive lists of attributes that a professional ought to display but little consensus on what professionalism actually is.

The level of professionalism displayed by a medic could be regarded as a perception. The expectation of the level of professionalism a physician ought to display is unique to the individual. With this in mind, one might wonder how there can ever be a standard that must be attained by medical students in order for them to be deemed professional.

The Oxford Dictionary[95] (p1141) defines professionalism as:

1. “The competence or skill expected of a professional.”

2. “The practising of an activity by professionals rather than amateurs.”

This definition may imply that professionalism is simply being competent and qualified. A dictionary definition of the construct of professionalism is low level. In reality professionalism is a complex concept, difficult to define, especially in a medical setting. It has multiple dimensions: ethical, moral and behavioural.

Van de Camp *et al* (2004) performed a systematic literature search in order to identify the constituent elements of professionalism mentioned in the definitions and descriptions of the concept. In all, 90 elements were identified, and are listed below. The elements most commonly mentioned were: altruism, accountability, respect and integrity [96].

Absence of impairment

Accountability

Adherence to guidelines

Altruism

Appreciate literature and arts

Ask help when necessary

Autonomy of professional associations

Avoiding misuse of power

Be knowledgeable

Be responsive to patients’ and colleagues’ age, gender and disabilities

Be sensitive

Be thoughtful

Being well-organized

Benevolence

Blow the whistle if necessary

Calling

Caring

Carry out professional responsibilities

Clear professional values

Commitment

Commitment to continuity of patient care

Communication skills

Compassion

Competence

Courage

Critical analysis

Critique

Deal with high levels of uncertainty

Deliverance of quality

Duty

Educate patients

Enhancing welfare of the community

Excellence

Expert authority

Faith in life’s meaning and value

Fight for and guarantee standards

Flexibility

Give patients information they understand

Good clinical judgement

Goodwill

High level of expertise

Honesty

Honour

Humanistic values

Humility

Integrity

Interpersonal skills

Justice

Know limits of professional competence

Leadership

Lifelong learning

Maturity

Method and thoroughness

Morality

Motivation

Negotiation

Not letting personal beliefs influence care

Not ripping people off

Participation

Understanding history

Use of explicit standards

Value medical work intrinsically

Virtue

Willingness to admit errors in judgment

Willingness

Professional awareness and sensitivity

Professional conduct

Protect confidential information

Relationships with colleagues/team

Reliability

Respect

Respect patients' right of shared decision making

Response to instruction

Response to stress

Responsibility

Self-awareness

Self-improvement

Self-regulation

Sensitivity to a diverse patient population

Service

Simplicity

Social contract

Submission to an ethical code

Suspension of self-interest

Technical competence

Temperance

Tolerance

Transparent rules

Treat patients politely

Trust

As well as the many authors, institutions and organisations who have provided their own definitions or attributes of professionalism , there are those as pointed out by Stern (2006) who, "claim they may not be able to define professionalism intelligibly but 'know it when they see it'[97]." As Stern comments, this may be an adequate stance for an individual, however the difficulty comes when groups must agree on what constitutes acceptable professional behaviour. In order to do this all members of the group must be in agreement on the same overall concept of professionalism, a task much harder than it seems, although this is perhaps self- evident if one looks above at the 90 statements currently in circulation.

In recent years there has been a substantial increase in the attention professionalism has been subject to across the medical curriculum. Its assessment is now at the forefront of medical education research. Unless a clear definition of professionalism is derived then it cannot be effectively measured.

A physician having to adhere to a professional code of conduct is not new. The Hippocratic Oath, estimated to have been written in the 4th century BC, was the code of conduct new physicians swore by, and reads as follows[98]:

I swear by Apollo the physician, and Asclepius, and Hygieia and Panacea and all the gods and goddesses as my witnesses, that, according to my ability and judgement, I will keep this Oath and this contract:

To hold him who taught me this art equally dear to me as my parents, to be a partner in life with him, and to fulfil his needs when required; to look upon his offspring as equals to my own siblings, and to teach them this art, if they shall wish to learn it, without fee or contract; and that by the set rules, lectures, and every other mode of instruction, I will impart a knowledge of the art to my own sons, and those of my teachers, and to students bound by this contract and having sworn this Oath to the law of medicine, but to no others.

I will use those dietary regimens which will benefit my patients according to my greatest ability and judgement, and I will do no harm or injustice to them.

I will not give a lethal drug to anyone if I am asked, nor will I advise such a plan; and similarly I will not give a woman a pessary to cause an abortion.

In purity and according to divine law will I carry out my life and my art.

I will not use the knife, even upon those suffering from stones, but I will leave this to those who are trained in this craft.

Into whatever homes I go, I will enter them for the benefit of the sick, avoiding any voluntary act of impropriety or corruption, including the seduction of women or men, whether they are free men or slaves.

Whatever I see or hear in the lives of my patients, whether in connection with my professional practice or not, which ought not to be spoken of outside, I will keep secret, as considering all such things to be private.

So long as I maintain this Oath faithfully and without corruption, may it be granted to me to partake of life fully and the practice of my art, gaining the respect of all men for all time. However, should I transgress this Oath and violate it, may the opposite be my fate.

If one is to very crudely summarise the Hippocratic oath, the following conclusion about the standards expected of physician could be drawn:

1. They must respect their instructors
2. They must pass on their knowledge to the next generation
3. They must do their patients no harm nor injustice
4. They will not perform euthanasia nor abortion
5. They will work within their limits/ capabilities
6. They must keep patient confidentiality

Issues such as euthanasia and abortion are controversial; their inclusion in a modern day code of conduct may not be plausible. This issue aside, the premise of the Hippocratic Oath resonates within the definitions of professionalism in use today.

1.12.1.1 The Declaration of Geneva, 1948.

The Declaration of Geneva was first adopted in September 1948 by the General Assembly at the World Medical Association, Geneva, Switzerland [99, 100].

This declaration was written as a revision of the Hippocratic Oath in response to the crimes committed by doctors in Nazi Germany which were prosecuted at the Nuremberg Trials [100, 101].

The declaration reads as follows:

At the time of being admitted as a Member of the medical profession

- I solemnly pledge myself to consecrate my life to the service of humanity :
- I will give to my teachers the respect and gratitude which is their due;
- I will practise my profession with conscience and dignity;
- The health and life of my patient will be my first consideration;
- I will respect the secrets which are confided in me;
- I will maintain by all means in my power, the honour and the noble traditions of the medical profession;
- My colleagues will be my brothers
- I will not permit considerations of religion, nationality, race, party politics nor social standing to intervene between my duty and my patient;
- I will maintain the utmost respect for human life, from the time of its conception, even under threat, I will not use my medical knowledge contrary to the laws of humanity;
- I make these promises solemnly, freely and upon my honour [99, 100].

The declaration has been amended in 1968, 1983, 1994 and 2005 [99]. The ancient Hippocratic Oath calls upon Apollo, Asclepius and the gods of healing[98], however, the declaration does not mention any form of god nor sacred power [101]. The declaration was made to reinforce the moral truths of the Hippocratic Oath, in a form that could be understood and accepted in the twentieth century. Similarly, the Declaration of Kuwait was

adopted at the 1st International Conference on Islamic Medicine, in Kuwait[102]. This was a specific code of conduct for Muslim Doctors.

The expectations of those within the medical field as to the attributes displayed or the standards attained by a professional are evident in the numerous papers written by faculty or guidelines[20, 103] published by governing bodies such as the General Medical Council (GMC) within the United Kingdom or the American Board of Internal Medicine (ABIM) in the United States of America. The attributes considered important within the profession, may differ anywhere from slightly to significantly from those that the patient might expect. Furthermore, as societal expectations of medical professionals change, so too will the accepted definitions of professionalism.

Not all definitions of professions are positive. George Bernard Shaw famously said that, "all professions are conspiracies against the laity", in The Doctor's Dilemma (1911) act 1. The alleged conspiracy could be demonstrated by the medical world within the UK, which is in fact self-regulated. Institutions working towards assessment of professional competence should aim to gain public trust by showing that unprofessional conduct will not be hidden from the laity.

Shaw's comments are best demonstrated by the Hippocratic Oath;

"...*To hold him who taught me this art equally dear to me as my parents, to be a partner in life with him, and to fulfil his needs when required; to look upon his offspring as equals to my own siblings, and to teach them this art, if they shall wish to learn it, without fee or contract; and that by the set rules, lectures, and every other mode of instruction, I will impart*

a knowledge of the art to my own sons, and those of my teachers, and to students bound by this contract and having sworn this Oath to the law of medicine, but to no others....”

Here the oath indicates that doctors should behave in a conspiratorial manner by offering the offspring of other doctors an in-group advantage. This evident tension between in- and out- groups in the Hippocratic Oath, one of the earliest professional records, may support Shaw’s remarks.

Today, many definitions of professionalism are in circulation. For the purpose of defining professionalism in the twenty-first century, I will use the definition offered by Stern;

“Professionalism is demonstrated through a foundation of clinical competence, communication skills, and ethical and legal understanding, upon which is built the aspiration and wise application of the principles of professionalism: excellence, humanism, accountability and altruism.”

1.13 Measuring professionalism

Professionalism lies at the heart of clinical practice [104]. Pressure from governing bodies, and more recently, patient awareness, has put increasing emphasis upon the role of a doctor as a professional.

In 1927 Francis Peabody commented that, " the most common criticism made at present by older practitioners is that young graduates have been taught a great deal about the mechanism of disease, but very little about the practice of medicine; or, to put it more bluntly, they are too "scientific" and do not know how to take care of patients" [105]. Almost a century later this criticism may still ring true. In order to address this issue, professionalism now rests high on the curriculum at medical schools worldwide.

Challenges not only exist for the academics endeavouring to assess professionalism in their students, but also for students who are confronted with mixed messages. Students are living in a world in which professional behaviour is being redefined, often in ways that run contrary to the medical education curriculum [106].

One such challenge faced by researchers is developing an assessment method that measures both cognitive and interpersonal skills, which are described by Dannefer *et al* (2005) as being of primary importance when considering professional competence.[107]

With professionalism being troublesome to define, particularly due to its subjective nature, assessment of professionalism can be equally as perplexing.

1.13.1 Can professionalism be assessed?

Good measures of cognitive knowledge and practical skills in undergraduate medical education are now available [108]. However, measures of professionalism are much more difficult to implement, stemming in part from the difficulty of defining professionalism, but also from measuring behavioural traits consistently.

Key problems are incompleteness of surveillance (students are not observed continuously), lack of accurate reporting (both peers and staff may not honestly report their perceptions for a variety of reasons), dissembling by students (who may role play expected behaviours) and the differences between teaching environments and real practice environments. However, it has been shown that disciplinary action through State Medical Boards is associated with concerns expressed during undergraduate training [109-111]. This suggests that the effort is well worth making, either in order to target remediation, or if that fails, to prevent progression.

Current approaches frequently rely on complex, subjective, and occasional decisions which are of low reliability, lack demonstrated validity, and are time consuming (hence expensive) to implement. They also have low defensibility, in that they are open to challenge by students. Any measure of professionalism, especially if being utilised as part an

assessment for student progression, must withstand legal challenges such as those presented in defence of a student at any potential Fitness to Practice proceedings.

1.14 Methods of assessing professionalism

There are numerous methods of assessing professionalism in both undergraduate and graduate medics currently in use. Regardless of the type of assessment used its value must be in its usefulness to future learning and its power to promote professional development.[107]

Existing assessment tools have been classified into three categories[112, 113];

1. Tools assessing professionalism as part of general clinical competence,
2. Tools assessing professionalism as a single construct,
3. Tools assessing separate elements of professionalism such as ‘humanism, self assessment, dutifulness, altruism, empathy and compassion, honesty, integrity, and ethical behaviour, as well as communication[113].’

Examples of professionalism assessments include reflection, self-assessment, workplace-based assessment, OSCEs, peer assessment, multi-source feedback and portfolios[97, 112].

Assessments of professionalism as part of clinical competence are often reliant upon rating scales and may be affected by poor inter-rater reliability[113, 114]. Objective Structured Clinical Examinations (OSCEs) and self-assessment are tools frequently described for assessment of one element of professionalism. However, OSCEs have been criticised for artificiality and self-assessments as unreliable in some settings[117].

Some of the methods utilised in the assessment of professionalism are now described in more detail. Particular emphasis has been placed on those most commonly used within undergraduate medical education; peer assessment, multi-source feedback and critical incident reporting. Portfolios and the mini-Clinical Evaluation Exercise are also described, however these are more commonly used during postgraduate training.

1.14.1 Peer assessment

Peer assessment is utilised as an assessment tool in chapters 6, 7 and 8 of this thesis. In light of the extensive discussion surrounding peer assessment and the experiment evidence supporting it provided in these chapters I will only provide an overview in this introductory chapter.

Peer assessment is the process of having members of a group judge the extent to which their fellow group members have exhibited specific traits, behaviours, or achievements.[115]

Peer assessment is a valuable source of information about medical students' professionalism [116]. By definition, peer assessment involves individuals who have attained the same general level of training or expertise, exercise no formal authority over each other, and share the same hierarchic status in an institution [117, 118] and is under used in medical education [119]. It can be a process that fosters reflection about both personal and professional qualities for many students [107].

Historically, peer assessment emerged in the 1920s. It drew attention during World War II as method to identify leaders among groups of servicemen [97]. Subsequently it was utilised by public service organisations under pressure to improve performance[97]. Peer assessment was first used in medical setting in 1950s as part of comprehensive assessment of performance rather than to assess specific professional behaviours.

Use within the medical setting typically results in a 'bifurcated description of competence: technical knowledge and skills versus nontechnical, interpersonal, humanistic qualities and skills[97].' Some of the examples of the humanistic qualities that peer assessment can evaluate given by Arnold and Stern include; integrity, responsibility, conscientiousness, empathy and respect[97]. Some forms of peer assessment result in "generalised descriptors of classmates' competence (e.g. suitability for a desirable residency). These not only offer insight into technical competence but also into principles of professionalism[120]."

Kane and Lawler (1978) suggested that the following three conditions are conducive to the use of peer assessment:

1) The existence of peer groups whose members are afforded unique views of salient aspects of each other's behaviour;
2) The existence of peer groups whose members are capable of accurately perceiving and interpreting the salient aspects of each other's behaviour;
3) A perceived need to improve the effectiveness with which some characteristics of peer group members are being assessed.[115, 120]

The use of peer assessment in assessing medical professionalism will be explored further in Chapter 7, however, I will describe specific types of peer assessment along with other measures of professionalism within this introduction.

One such method is peer nomination, in which students are asked to nominate a limited number of classmates who best fit various characteristics [115, 121] . This method was demonstrated by Arnold *et al*, 2005 [117] and is the most frequently researched method of peer assessment [115], an example being where students are asked to identify the top three peers that they would like to have work at their side in a medical emergency [116]. The student making the nominations usually has to exclude themselves from those they consider when making their nominations [115] .

The advantages of using peer nominations are that this method effectively discriminates extremes [115] and it does so with a high degree of validity and reliability. This method is

“virtually useless” as a feedback device [115, 120] as it “furnishes no explanatory information at all about the typically substantial proportion of group members who receive no nominations[115].”

Peer rating is a method where students are asked to rate peers for a series of questions or characteristics according to a fixed scale [115, 120, 121]. The scales most frequently used in the rankings are behaviourally anchored rating scales [115, 121]. An example of this method is where a student is asked to rate how frequently a peer demonstrates clinical competence using a scale ranging from ‘always’ to ‘never’ [120]. Scoring is simple and is done by averaging the ratings [115]. This method is most commonly used in medical education, as it is easy to administer and score. It provides specific information which is valuable for feedback. However, it is less reliable and valid [120]. Peer ratings should be used when “unique descriptive information is desired about each ratee[115].” This method is less valid and reliable than peer nominations [115, 120].

The third and final method of peer assessment is peer rankings. This method is the least researched [115]. Peer rankings require students to rank every peer from best to worst on one or more factors[120, 121]. This is the method of choice when trying to discriminate between all students [115]. It discriminates between the entire range of performance levels better than the other methods described [115]. It is thought that the validity and reliability of peer rankings are acceptable [115, 120].

The strengths of peer assessment are evident, and are highlighted by authors as discussed below.

1.14.1.1 Peer assessment: the positive aspects

Medical student peer evaluation has been suggested as a better predictor of performance during residency training than medical school grades or faculty evaluations and it seems to be stable from year to year. [119] Peer assessment can be beneficial or detrimental depending upon how it is carried out [118]. How best to facilitate peer assessment of students' professional behaviour remains to be answered.[116]

Dannefer and colleagues (2005) suggest that peer assessment holds promise for evaluating technical and interpersonal skills, but it "seems uniquely suited to the latter.[107]" The authors proceed to comment that, "peer assessment can provide a measure of difficult-to-access interpersonal skills at all levels of medical training and that medical students are capable of giving and receiving constructive feedback[107]."

Other research has deduced that it is possible to obtain assessments from associates in areas such as clinical skills, humanistic qualities and communication skills.[114] Ramsey *et al* concur with Dannefer, that peer ratings provide a practical method to assess clinical performance in areas such as humanistic qualities and communication skills that are difficult to assess with other measures [114, 122].

Arnold and Stern suggest that teachers spend less time with students than previously was the case; therefore they do not have the opportunities to observe students displaying a wide range of professional and unprofessional behaviours. This is more likely to occur among peers as they spend more time with each other and work closely in teams. Furthermore, it is the "non-hierarchical relationships among peers that promote authentic behaviour and genuine feedback, while reducing the biasing influence of social desirability[97]."

There are, as with all assessment methods, difficulties associated with the use of peer assessment.

1.14.1.2 Peer assessment: the negative aspects

Reluctance and refusal to participate have been noted [117, 122]. Potential risks associated with peer assessment have been reported and include students making inappropriate and harmful comments, and others having difficulty interpreting and making use of their feedback[107].

Arnold and colleagues conducted focus groups in order to ascertain the reasons for students' reluctance to participate in peer assessment[117]. The authors noted that the students interviewed grappled with the prospect of looking bad themselves if they brought a peer's unprofessional behaviour to the attention of someone; they were concerned for the well-being of peers. Some students reported that they wanted to give peers the benefit of the doubt, particularly in ambiguous situations as they were unsure if they were interpreting

the behaviour of a peer correctly. Concerns were expressed that there may be negative consequences for a peer as a result of a negative report. Additionally, students cited a reluctance to damage personal relationships by hurting their peer's feelings or incurring a peer's anger. Furthermore, they did not wish to disrupt relationships among teammates nor members of their work groups[117].

Arnold and Stern identified many factors which influence students' willingness to participate [117]. These factors can be assigned into several domains. Some of the factors that their students' identified as influencing their decision to participate in peer assessment have been summarised in table 1.

Domain	Student comments
Content	• Students should evaluate specific behaviours (areas in which they have expertise) • Positive and negative behaviours ➢ some students think a focus upon negative behaviour would do little to promote peer assessment
Anonymity	• Anonymity encourages genuine participation in peer evaluation • Protects peer & evaluator • Concerns that total anonymity gives students opportunity to vent or have a vendetta and not be accountable
Who	• Not comfortable assessing more advanced student • More comfortable assessing junior student • Concern over who receives the peer assessment once completed
Use	• Preference for strictly formative use • No consequences/ some thought there should be consequences for peers receiving negative feedback

Table 1: Factors influencing students' willingness to engage in peer assessment.

Analysis of student questionnaires by Van Rosendaal *et al* [122] revealed student opinion that any form of peer evaluation was seen to be of little benefit to the students individually and a major intrusion into the relationships existing among the residents questioned. There was also an anticipation that a substantial undermining of working and personal

interrelationships would result from the process[122]. Relationships between students are of high importance in terms of their ‘education and psychosocial experience’. Van Rosendaal and colleagues speculated that the need to defend this may prove to be an impediment to the inclusion of peer assessment in residency evaluations[122]. This is an issue institutions implementing any form of peer assessment must consider: will relationships compromise the efficacy of the assessment?

There are a number of competencies that can be judged by peer assessment [118]. With this in mind it is of no surprise that generalisation is difficult. There are multiple factors, including reliability, relationships, stakes and equivalence that may influence the quality of peer assessment[118]. Peer assessment is still seen by some as unreliable, and as “little more than personal recommendations from friends[114].”

Contributing factors to the reliability of peer assessments are the number of relevant performances observed, the number of peers involved and the number of aspects of competence being evaluated [118]. Performance of an individual varies between situations, and therefore, particularly in a clinical setting, reliable results can only be obtained from a number of observations. This is also true of the undergraduate setting, observations should be made over a period, rather than being based solely upon a single event.

Experienced evaluators differ when observing the same event [118]. In order to overcome this, multiple evaluators are needed to attain reliable results. Estimates suggest that 11 peers are required to achieve a reliability coefficient of 0.70 [114].Use of 11 peers can

identify physicians who are significantly above or below their peers[114]. With specific regard to peer assessment, additional peer evaluators often contribute different observations [118].

In addition, multiple aspects of competence must be assessed [114]. Ramsey *et al* found that 10 questions regarding different aspects of competence can reliably assess competence [114].

As mentioned previously, relationships can effect the reliability of peer assessment [118]. However this can usually be overcome by making all evaluations anonymous.

1.14.1.3 Implementing peer assessment

Norcini (2003) lists important steps that ought to be implemented when conducting a peer assessment. The assessment purpose should be stated, preferably in writing, and ought to be communicated to participants. Participants should be aware of what is expected of them as both the evaluators and the subjects[118]. Peer assessment can be introduced gradually, starting with anonymous evaluations in a low stake setting[118]. It has been suggested that peer assessment should only be used for formative purposes[97].

Assessment criteria must be developed and communicated to participants, and training must be given. The results of the assessments should be monitored throughout the process, this can include checking reliability of results and eliciting feedback from

participants, which will allow for amendments to be made as necessary. Norcini's final recommendation is that feedback should be given to participants, appropriate to the purpose of the assessment[118].

Arnold *et al* proposed a format for peer assessment that could be implemented across all medical schools. This was derived after questioning students at multiple medical schools in order to ascertain the criteria that they thought such an assessment should meet[116]. The suggested model is as follows:

The peer assessment would clearly be 100% anonymous and would be designed so that the relationship between the student evaluator and the classmate would not be compromised. It would be a part of all clinical rotations and occur immediately after the classmate behaved either unprofessionally or professionally. Participating students would be asked to assess both positive and negative professional behaviours of peers who were at the same year level. In addition, each student would receive an annual summary of strengths and weaknesses identified by peers [116].

As an adjunct to this the authors proposed that institutional support ought to be offered to the students. The support available should include the faculty clearly defining what expected professional behaviour is. Training in how to give and receive peer feedback would be offered to students. Faculty and residents would evaluate their peers and thereby offer role modelling of peer assessment. Most importantly, the school would take

appropriate action regarding the behaviour of students in light of information it received about the behaviour from peers[116].

In order for peer assessment to be implemented successfully McCormack suggests that the assessment must be administered in the least threatening environment possible [115, 120]. Also, students' motivation is driven by their understanding of the benefits they will receive from participating and these benefits should be made clear. Examples given by McCormack included long-term feedback and possibly more importantly, short-term feedback such as Dean's letters[115, 120]. For the optimum result from the peer assessment process students must know their peers well enough. Students should be given the opportunity to select the peers they wish to evaluate, or be given an "unable to judge" option on rating scales. Finally, students should not be asked to evaluate anything that they have not seen first-hand; an example given by McCormack was first year medical students who may be able to evaluate communication skills but not clinical skills[115, 120].

1.15 Multi-source feedback

Multi-source feedback (MSF) is not contained within the experimental chapters of this thesis. However, it is included within this introduction as it is one of the most popular mechanisms by which professionalism is assessed, especially within the clinical setting. Issues concerning MSF such as validity, reliability, delivery and acceptability are also transferable, to a certain extent, to the larger debate about assessment of professionalism.

Multi-source feedback, often also known as 360 degree feedback is defined as: the systematic collection of data and feedback of performance data on an individual or group, derived from a number of the stakeholders in their performance [123].

There is debate as to whether the terms MSF and 360° feedback are interchangeable [124]. According to Foster and Law, multi-source feedback requires only 2 raters, whereas 360° feedback requires a full circle of feedback. The argument used as to why MSF and 360° feedback are not interchangeable is that a self-rating and a peer-rating on the same individual constitute multi-source, however this is not a 360 degree picture. The authors produced a diagram, as outlined below to describe this circle of feedback.

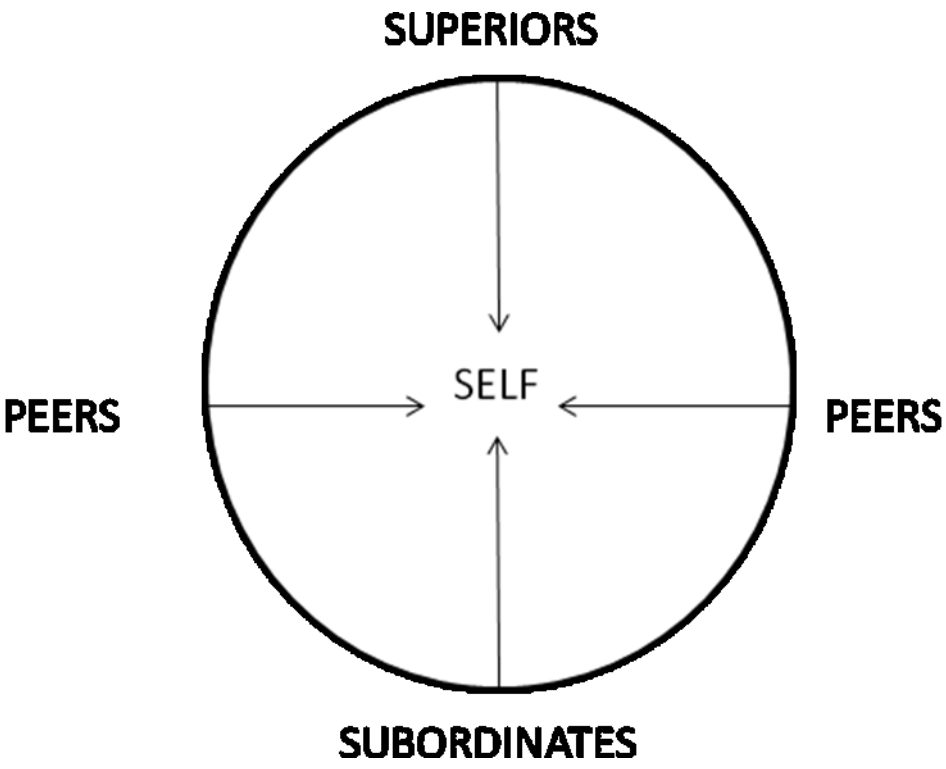

Figure 2: The 360° feedback model. (Adapted from Foster and Law, 2006[124])

Foster and Law state that the circle can be created using members of “organisational hierarchy.” The model demonstrates how superior-ratings, peer-ratings and subordinate-ratings can be conceptualised in the form of a circle around an individual who provides self-ratings. The authors suggest that 360^{o} feedback is a particular type of MSF, due to its basis upon two important organisational principles, proximity and power.

For the purposes of this introduction 360 degree feedback will be considered to be a specific type of multi-source feedback, which utilises self, superior and subordinate raters. Multi-source feedback will describe any feedback which utilises two or more raters, can include any number of raters and does not always have to include self, superior and subordinate raters.

Within a medical setting the MSF uses assessments from patients, medical colleagues, co-workers and self for the assessment of doctors [125, 126].

When “used effectively MSF has the potential to generate structured feedback, which informs educational planning by building on strengths and indentifying areas for development[125].”

1.15.1 The structure of MSF

It has become evident from literature that MSF has 4 stages, whether it is being utilised in a medical setting or another industry. These stages are:

1) Survey design
2) Role set selection (the stakeholders mentioned previously are selected to provide feedback on practice)
3) Data collection
4) Score calculation and feedback

1.15.2 Survey design

The questionnaires used usually consist of a number of statements on a rated scale. There is often the opportunity for accompanying free text comments [127]. Research has found that the traditional survey method of using a paper-based pro forma yields better results than internet or phone methods [128].

Considerations to be made when designing the assessment are that the completion of the assessment should be 'convenient and appropriate' to stakeholders and manageable those sorting the data (i.e. not too data to distribute or analyse) [126]. Those being assessed should be consulted when designing the questionnaire [127]. Questions must be designed to ensure that the raters are judging what they perceive as behaviour and not the intentions behind it [123].

1.15.3 Role set selection

Within a medical setting the role set commonly includes patients, medical colleagues, co-workers and self for the assessment of doctors [125, 126]. These co-workers can be of differing grades to the assessee e.g. junior doctors and consultants or can be inter-professional e.g. nurses, administrative or other allied health professionals. Increased numbers of raters will allow wider sampling both across and within occupational groups broadening the perspective [125]. Assessees who, after analysis of results, appear to be borderline may require more raters [125].

With specific reference to the 360 degree model, it has been suggested that when too few raters are recruited it can be troublesome because this undermines the principles of proximity and power that make 360 degree feedback distinctive [124]. Omission of superiors, peer or subordinates removes part of the circle (defined above) necessary for the full picture. It has been argued that if any category of rater was to be omitted then it could be the self-rating. Although the authors did proceed to conclude that self-ratings are necessary to evaluate the discrepancies between self and others, and therefore “provide the hub of the 360^{o} circle” [124].

1.15.4 Data collection

There appears to be debate as to whether or not the raters should remain anonymous to the assessee. Garbett concluded that, “ anonymity is not a prerequisite to the provision of open, honest feedback” [126]. Data collection can either be paper-based or electronic, the latter is advantageous as it is automated and makes it easier to track [123].

1.15.5 Score calculation and feedback

Analysis will vary depending upon the type of questions asked, and whether they use scales or free-text for responses. Analysis of themes arising from qualitative data can be done in collaboration between the rater and the assessee [126].

Ideally feedback should be face-to-face [125]. Those being assessed should be allowed to read through the feedback before the face-to-face discussion. During the discussion, the focus should be upon identifying the areas of strength and weakness and helping the learner to identify where they feel that developmental work would be useful [125].

Where feedback is given as a report, it should summarise the answers given, showing actual ratings given for each question, as well as averages for each questions. It should also include any written comments [127].

Ward notes that even though the feedback received comes from people the assessee knows it does not eliminate the chance of it being “quite pointed, surprising” and at a variance with their own view of themselves. The feedback facilitator therefore has to help the assessee to “accept and internalise the data before they can act on it” [123].

1.15.6 Use of MSF in medical education

According to a survey by Stephenson, in 2005 only 5 of the 29 UK medical schools were formally using MSF [129]. The 42 students from across the 5 schools who received MSF were questioned about their understanding of MSF. Only 1 student had heard of MSF and was able to summarise the process. Other students had "a degree of confusion ... regarding 360° [127] and their assessment procedures."

Stephenson[130] explained MSF to all 42 students and canvassed their opinions. The comments were widespread, and included concerns over staff willingness to evaluate and appreciation for the opportunity to be assessed by a range of colleagues. Examples included:

1. "Don't need the extra hassle of running after more staff to assess me."
2. "Would it really work? Medical staff are always busy, would people really take the time to assess us?"
3. "… time intensive and requires lots of staff…"
4. "Aren't we assessed enough?"
5. "Being assessed by junior doctors would be good because we spend a lot of time with them."
6. "Different professions will look for different qualities and therefore will be able to provide better feedback."
7. "I spend more time with nurses so it would be nice if they could have a say in my assessment."

8. "… It would show a multi-professional approach."
9. "Consultants should not have the only say in our assessments."
10. "More rounded view."

The above comments presented by Stephenson[130] demonstrate issues as perceived by stakeholders in professionalism assessments, it is for this reason that I have included them within this introduction. Any future professionalism measure should bear these issues in mind as they will impact upon the validity and acceptance of any proposed interventions.

1.15.7 Concerns about assessment of professionalism using MSF

Doctors are concerned about threats to the self-regulating nature of the profession. Unfortunately, the assessment of doctors, especially using MSF, can be perceived as threatening or intrusive, rather than as providing opportunities to define individual strengths and identify areas of learning that need to be developed [131]. Feedback from multiple sources simultaneously could be a cause of anxiety.

1.15.8 Benefits of using MSF

If completed to a high standard MSF can challenge the perceptions of the assessee with respect to their skills and performance. It can provide them with the motivation to change. [127] The Institute of Personnel and Development suggest that MSF challenges the assessees' perceptions in 3 ways:

1. The feedback on an aspect of behaviours is the opposite of what the recipient expects

2. An aspect of the behaviour is shown to be more (or less) important as an aspect of their performance than the recipient thought.
3. The results highlight relationships between the aspects of behaviours.

“Critical feedback provides the motivation to change, as long as the raters are credible and their views are of value to the assessee” [127].

It is not only the individual being assessed that stands to gain from the MSF process. Garbett suggests that teams within an organisation can benefit from developing cultures where there are frameworks in place that allow people to offer feedback on each others’ performance [126]. The active seeking of feedback from a wide range of sources “exemplifies responsiveness in an era when organisations change rapidly and service users expect action in response to their comments” [123], a comment certainly true of the National Health Service and its patients (users). Within a medical setting Garbett suggests that MSF is useful to practitioners seeking to understand their practice.

MSF can be useful for doctors who lack confidence [125]. It offers doctors a valuable opportunity to obtain structured feedback from a range of colleagues that helps them to obtain a better understanding of their role and functioning as a professional, facilitating meaningful professional development [125]. MSF also helps participants to develop more open relationships with colleagues and service users [126].

MSF benefits employers (or in this case, faculty) as it provides “new pertinent information about competencies” [127] and how others see individuals. It can also be valuable for

assessing or implementing new initiatives such as training, development and appraisal [127].

1.15.9 Problems associated with MSF

Lockyer reported that the mobility of doctors was a problem when carrying out MSF and suggested that industry guidelines should be followed. Such guidelines suggest that MSF should be carried out after a period of 6-12 months in a given workplace.[128] Mobility is not an issue during the pre-clinical years of undergraduate curriculum but could potentially be problematic during clinical training and post qualifying where students rotate on a frequent basis. This might make it difficult to implement MSF for a variety of reasons. A major concern could be that students might not establish relationships with or work with the same colleagues enough (if working shifts) for them to be in a position to complete an evaluation. This could suggest that the pre-clinical years of their undergraduate education is a good time to instigate MSF.

The Institute of Personnel and Development has identified many concerns that employees may have with regard to MSF. In order to eliminate as many of these as possible they suggest the following guidelines are adhered to [127]:

- Issues of confidentiality are clearly communicated detailing who will have access to the data and for what purpose.
- Clearly stated how feedback will be given and by whom.
- The process for identifying respondents is clearly set out with recipients having some opportunity to input.

- Sufficient time is allowed to pilot the process and to consult with individuals and employee groups on both the design and implementation of the process.
- Both recipients and respondents are adequately briefed on the process, how to complete the relevant forms and the aims and objectives of the exercise.
- Adequate opportunity is given for people to comment and raise their concerns.
- People are not forced nor coerced to take part by managers.
- Feedback is never attributed to an individual, that feedback reports and developments plans are kept secure and that data protection rules are obeyed
- The process is constantly monitored and evaluated, all concerns are acted on and any changes are adequately communicated.

Although written with a different industry in mind, these guidelines can be easily adopted by faculty wishing to implement MSF into their undergraduate curriculum. Adaptations of the above guidelines may be necessary if MSF was being utilised as a formal assessment method, it may become necessary to make completion of evaluations compulsory.

1.16 Critical Incident Reporting

Critical incident reporting is an evaluation in context, which provides detailed descriptions of specific sentinel events. These retrospective assessments of behaviour are useful for students at the extremes of performance.

Critical incident reporting is a method of reporting both professional and unprofessional behaviour. It is utilised at Durham University but is known as Significant Event Recording in

order to capture both positive and negative events. Chapter 9 analyses the use of Critical Incident Reporting within the undergraduate medical setting. This introduction serves to highlight the use of critical incident reporting as a tool for documenting (un)professional conduct.

The structure of Critical Incident Reporting will differ between institutions, however, the premise remains the same. When an event of note occurs, be it positive or negative, a Critical Incident Report form is completed. This form documents the sequence of events, and an explanation of the impact of the event and often provides chance for the recipient to reflect upon the event. Critical incident reporting is used to characterise individual learners [113].

In previous studies[132-134] analysis of critical incident forms identified residents with unprofessional behaviours. Rhoton also described many types of unprofessional behaviours. Those most frequently identified included expressions of personality problems, fabrication and abdication of responsibility. Predictors of unprofessional behaviour were also identified, including deficiencies in conscientiousness, taking instructions, eagerness to learn and efficiency. Rhoton's final finding was that residents with no instance of unprofessional behaviour in their records achieved excellent clinical performances, where as those who with critical incident reports filed against them performed poorly.

Arnold suggests that future research on the use of critical incidents should include the reporting of less severe events as well as the major lapses in professional behaviour. This would enable patterns of unprofessional behaviour to be identified [113].

Other studies evaluating critical incident reporting have performed longitudinal assessments which track students' professionalism throughout medical school, with the goal of offering remediation where necessary. Arnold explains that this approach allows faculty to "quantify their impressions of problematic students in a uniform manner" [113]. The form Arnold describes lists behavioural indicators of traits of (un)professional behaviour, although this would inevitably vary between institutions. Generally the faculty form reporting the unprofessional behaviour (inside or outside of class) is forwarded to the Dean, who will meet with the student to decide on the appropriate action to be taken. The most common incidents reported relate to lack of conscientiousness or poor relationships with the health care team [113].

1.16.1 Problems associated with the use of critical incident reporting

As the focus of critical incident reporting tends to be on unprofessional behaviour, or in some institutions both positive and negative extremes of professionalism, not all students receive feedback. Only those students who have a report filed will receive feedback from the Dean [113]. Students receiving differing amounts of feedback depending on their behaviour may be considered unethical by some, and therefore if critical incidents were to be implemented by a faculty, this is an issue that they would have to address with staff and students.

1.16.2 Advantages associated with use of critical incident reporting

Hodges *et al* suggested that when students are asked to reflect upon the critical incident in question it provides them with the opportunity to examine their behaviour and attitudes, to consider why the incident occurred and to allow solutions to be developed that could help prevent such incidents in future [135]. This is the study presented in Chapter 9 of this thesis. These findings are in accordance with Papadakis and Loeser's statement that critical incident reporting triggers transformative learning through reflection on behaviours[136]. This transformative learning is specifically in relation to reflection on the underlying reasons for and assumptions about the behaviour exhibited.

Another advantage of the system of critical incident reporting is that it allows for the possibility that single events do not necessarily define professionalism. This is consistent with the 'definition of professionalism as a set of behaviours, rather than a dichotomous dimension'[136], in which every professional exhibits a range of behaviour including lapses.

Critical incident forms are a cost effective method of recording and tracking student behaviour. If stored, they will allow future tracking and multiple forms could potentially identify students at risk of failing or exhibiting severe unprofessional behaviour. This will allow remediation to be offered.

1.16.3 Evidence supporting the use of critical incident reporting

Case Western Reserve Medical School, Ohio, devised the Clinical Anesthesia System of Evaluation (CASE) to track any critical incidents associated with residents. Examination of the 'CASE' provides an illustration of the sorts of issues critical incident reporting can be used for within the clinical setting.

When comments from the attending physicians supervising residents were analysed, the critical incidents fell into the following categories:

1. Omissions that would or could have caused significant morbidity or mortality had teachers not intervened (e.g. failure to recognise an obstructed airway)

2. Unprofessional behaviour:

    -Failure to follow specific orders

    -Falsifying reports regarding activities or therapy

Famously, a case-control study was conducted at the University of California, San Francisco, School of Medicine (UCSF) where Papadakis *et al* [110] considered all UCSF medicine graduates who had been disciplined by the Medical Board of California between 1990-2000 of which there were 68, and compared them to a control group of 196 graduates.

The main reason for the disciplinary action was for a violation in the physician's professionalism (65 out of the 68 cases). Disciplined physicians were more than twice as likely to have negative comments regarding professionalism in their medical school files (odds ratio 2.15, p=0.02)

Traditional measures of medical school performance, including grades and passing scores on national standardised tests, did not identify students who had disciplinary problems as practising physicians.

The paper demonstrated that, "unprofessional behaviour in medical school is associated with unprofessional behaviour in practice." The authors suggest that this finding can now serve as evidence to some students that they must commit to professional growth.

In another case-control study Papdakis *et al* [109] again found that disciplinary action by a medical board was strongly linked to previous unprofessional behaviour in medical school (odds ratio 3.0). The types of behaviour most strongly linked with disciplinary action were severe irresponsibility (odds ratio 8.5) and severely diminished capacity for self-improvement (odds ratio 3.1). Examples of the irresponsible behaviours documented were unreliable attendance at clinic and not following up on activities related to patient care.

Those with a diminished capacity for self-improvement had failed to accept constructive criticism, displayed poor attitude or were argumentative. There were further links associated with disciplinary action, although they were not as strong as the links between

unprofessional behaviour and disciplinary action. These included low scores on the Medical College Admission Test and poor grades in the first two years at medical school.

The conclusion that can be drawn from both of the Papadakis studies is that unprofessional behaviour as a medical student is the strongest predictor of disciplinary action.

Studies by Papadakis have provoked a number of papers within Medical Education examining professionalism. Her finding that poor behaviour as an undergraduate can be predictive of unprofessional behaviour as a practising physician has led many, including myself, to seek measures of professionalism which can be utilised during training to track and highlight those students who could potentially face future Fitness to Practice proceedings.

1.17 The mini-Clinical Evaluation Exercise

The mini-clinical examination (mini-CEX) [137], also known as the mini-clinical evaluation exercise[138], is a method of assessing competency during clinical education. A mini-CEX can be utilised to evaluate history taking, clinical examination, communication skills, clinical judgement and professionalism[137].

1.17.1 Aims of the mini-CEX

These exercises strive to assess trainees in a holistic manner, considering professionalism alongside technical competencies and communication skills[112].

As with other forms of workplace-based assessment, the primary purpose of the mini-CEX is to provide structured teaching and feedback in relation to clinical practice[138]. The assessment and feedback is simultaneous[139]. Assessments form part of a trainee's portfolio and can be gathered in a variety of settings[138]. Mini-CEX can be utilised within undergraduate training, as well as with postgraduates (p195)[112].

1.17.2 Advantages of the mini-CEX

A reported strength of the mini-CEX is that it adapts itself to a broad range of clinical scenarios and situations[140]. Norcini presented data to show that examiners' ratings did not differ substantially depending upon which training programme the mini-CEX or the examiner was associated with, nor did they differ depending upon the nature of the patient[140]. Norcini's findings show that the mini-CEX is advantageous as it is possible to compare ratings based upon a variety of clinical problems. Norcini states that this makes the mini-CEX a valuable tool for use at the local programme level[140].

Hill and colleagues have also shown the mini-CEX to have good overall utility for assessing aspects of clinical practice within the undergraduate setting[141]. They report the strengths of the tool to be its fidelity, wide sampling, perceived validity and formative observations and feedback. In accordance with other literary evidence [112, 139, 140, 142, 143], Hill *et al*

suggest that the mini-CEX has limited reliability dependent upon variations in examiner stringency[141]. Unfortunately in a randomised control trial, Cook *et al* found that rater training did not improve interrater reliability nor accuracy of scores[144]. Confounding variables which may affect the validity of the mini-CEX are reported as examiner status, the complexity of the case being assessed, the focus of the case, and patient gender[141]. Some of the confounding variables, such as impact of the patient and the case being assessed, listed by Hill *et al* [141] are contrary to the findings of Norcini[140] described earlier. In light of this conflicting evidence, institutions should exercise caution if employing the mini-CEX as their only tool for assessing professionalism.

1.17.3 Disadvantages associated with the mini-CEX

As with all assessment methods, the mini-CEX can be affected by role conflict, stakes and equivalence [142] (p150). Avoiding role conflict is troublesome; however to minimise impact Norcini suggests using faculty observers who are not directly involved in the teaching of those being assessed. The rationale for this is that faculty are removed from assessing the outcomes of their own educational efforts.

The stakes of the assessment could also impact upon the reliability of the mini-CEX. Faculty observations may not be as accurate depending upon their use. This is especially true if the consequences of an assessment outcome are adverse to the student. Norcini[142] suggests that maintaining the anonymity of the faculty assessor may reduce the influence of stakes on the assessment outcome.

Finally, equivalence may also impact upon the reliability of the mini-CEX[142]. As the activities being undertaken by the student during an evaluation may differ, scores for different trainees may not be comparable. Additionally, the stringency of assessing faculty members will differ. In order to account for a potential lack of equivalence it is suggested that the number of assessing faculty is increased and that evaluations should take place across a broad range of activities. Furthermore, faculty training may reduce the impact of non-equivalence in marking stringency.

1.17.4 Evidence supporting the use of the mini-CEX

Hill and colleagues demonstrated reliability coefficients of 0.73 when 15 encounters were aggregated. However, they suggest that the mini-CEX might not be valid as a tool for assessing specific, individual competency domains. Whilst the authors advocate the use of the tool in undergraduate assessment they call for its use to be "tailored to fit the overall assessment strategy of each institution[141]."

Scores are recorded using Likert scales[112, 139, 142]. Nine-point Likert scales have been shown to provide more accurate scores than five point scales. However, inter-rater reliability is similar for both nine and five point scales[145].

1.17.5 Use of the mini-CEX

Mini-CEX have been described as an assessment tool for recording professional conduct[97] (p148). Examples of professional conduct which the mini-CEX is commonly

used to assess are the student shows respect, courtesy, empathy and establishes trust, the student attends to the patient's needs of comfort, respect and confidentiality, and the student behaves in an ethical manner with awareness of legal frameworks and of their limitations[139]. Use of the mini-CEX as a professionalism measure has been categorised as 'faculty observations' by Norcini[142]. Norcini describes how these 'single encounter' judgements are advantageous as they 'focus faulty on a particular performance' and ensure that students are observed. As performance during one observation is not predictive of future performance, several observations must be made to ensure reliability. These individual assessments would provide the overall evaluation by aggregation of the scores[97].

When utilising the mini-CEX as assessment tool for professionalism, humanism and professionalism are often recording using Likert scales[142]. Evidence presented by Norcini *et al*[143] and Kogan *et al[146]* reported that the highest marks given for a mini-CEX were for scores of professionalism. Furthermore, as the students progressed through their training, their professionalism scores increased. With respect to students who were borderline, at least 4 observations were deemed necessary to provide results in which faculty could have confidence[143].

The mini-CEX provides an assessment of professionalism which assesses "whether students know and can show how to demonstrate professional behaviour (p156) [142]." The success of the mini-CEX is dependent upon its purpose being specified, adequate development of the assessment criteria, faculty training, informing students , monitoring the programme and providing routine feedback[142]. The validity and reliability of the mini-CEX

is derived from ‘the fact that trainees are observed while engaging with a series of real patients in different practice settings and judgements about the quality of those encounters are made by skilled educator-clinicians’[139]. Importantly, the mini-CEX is highly regarded by faculty and trainees, imperative to successful implementation[141].

1.18 Objective Structured Clinical Examinations

Objective Structured Clinical Examinations (OSCEs) have been described as an assessment method for professionalism [112, 117, 147]. These examinations, primarily utilised for assessing a candidate’s clinical examination and communication skills, present scenarios which are usually delivered by a simulated patient. In recent years OSCEs have been employed for assessing professional conduct.

Use of the OSCE as a tool for assessing professionalism has been criticised [97, 113] for being artificial, cost prohibitive to most institutions and for having low reliability. The latter is a result of needing multiple scenarios in order to accurately assess a candidate under hypothetical conditions, such as those within the OSCE. Stern[97] (p49) suggests that the use of OSCEs for assessing professionalism should be for formative purposes only.

Durham University does not utilise OSCEs in summative professionalism assessments. Examiners are permitted to indicate a global judgement regarding students’ professionalism on their score sheets. This is separate to the marking criteria for the clinical skill being examined. Global judgements enable faculty to record concerns regarding a student’s

professionalism when the candidate may otherwise be performing well on the clinical task being examined.

1.19 Portfolios

Portfolios are another means by which a student's professionalism can be assessed. "A portfolio is a purposeful collection of evidence gathered by individuals on their roles as learners, recording and reflecting on a learner's progress and achievement in selected domains[148]." Such portfolios can become tools for reflection[148], making them a useful means by which professional conduct can be recorded and reflected upon.

It has been argued that due to the integrated system of learner-directed evaluation which constitutes a portfolio, portfolios are well suited for both formative and summative use in the assessment of 'complex and multifaceted skills and competencies[148]', including professionalism. However, Roberts *et al* (2006)[149] expressed concerns about the reliability of portfolios in high-stakes situations.

Portfolios, the use of which originated in subjects such as art and architecture as a platform for logging development, have become popular in medical education. Thistlethwaite and Spencer[112] attribute this to their congruence with the 'nature of professional practice and how professionals learn (p196).'

Portfolios are a means by which assessments can be organised, however, whether they are an assessment tool is a contentious issue[112].

1.19.1 Advantages associated with the use of portfolios

Portfolios are advantageous as there are no limits to their content. Examples of content provided in the literature[112, 148] include peer and self assessment, case reports, checklists and evaluations, to name but a few. Another advantage of portfolios is the links they create between learning and assessment[112].

The use of portfolios in assessment of professionalism is associated with their flexibility, as traditional assessment such as multiple choice questions are constrained. The portfolio enables exploration of attitudes, attributes and reflection[112]. Generally, a portfolio is completed over time, lending it to assessment of professionalism as this encompasses the developmental nature of professionalism.

Thisthlethwaite and Spencer provide an exhaustive list of the perceived advantages of using portfolios in assessment[112]. These include assessing areas which are difficult to test conventionally, encapsulating principles of adult learning, they incorporate a range of quantitative and qualitative methods, they assess authentic and workplaced-based contexts, they promote the link between theory and practice and they stimulate reflection. Furthermore, portfolios stimulate creativity, they provide a development of strategies and life-long learning and profile the learner's competencies and achievements. Interactions

between learners and their teachers are enhanced, and the portfolio can take account of the learning style preferences of individual students[112].

1.19.2 Disadvantages associated with the use of portfolios

As with all assessment tools, portfolios have their draw backs. These are noted to be concerns about reliability, the possibility for plagiarism and fabrication, time constraints for both the trainee and the assessor and costs[112]. Negative attitudes towards portfolios are also reported[112], acceptability is important when considering implementation of any assessment method[150].

Finally, considerations for institutions contemplating the introduction of portfolios as an assessment tool include defining a clear purpose and content, provision of instructions for use, transparent marking criteria and a full evaluation of the process[112]. The format of the portfolio is also an important consideration as this can be paper-based, web-based or video[148].

1.19.3 Evidence supporting the use of portfolios

Fryer-Edwards *et al*[148] describe how use of a portfolio can engage students in deep approaches to learning, resulting in students being able to perform ‘capably and professionally in varied environments (p217).’ Portfolios are effective vehicles by which learning can be goal oriented, student centred and experiential[148]. This is in part due to the reflective nature of a portfolio. Reflection translates the experience of clinical practice

into learning[148]', and perhaps is partially due to the broad range of learning activities it encompasses.

McMullan *et al* (2003) suggest that a holistic approach to competence appears to be compatible with the use of portfolios to assess competence. They advocate the use of a variety of methods when assessing competence but suggest that portfolios have the potential to integrate these methods[151].

Reflection is an essential component of a portfolio, as are the student–teacher relationship and explicit guidelines for constructing the portfolio. Issues of rigour in assessment of portfolios need to be addressed, but the assessor's professional judgement will inevitably enter into this assessment[151]."The portfolio is much more than a document providing only evidence of what has gone before: it is a dynamic record of growth and professional change[151]." It is this quote that summarises why many institutions are choosing to adopt portfolios as part of their assessment of professionalism.

1.20 Conclusions: Professionalism

Lynch and colleagues summarise the efforts of many to assess professionalism; "although assessing professionalism poses many challenges, gauging and ascertaining growth in professionalism is impossible without assessment[152]." In order to compensate for weaknesses in current assessment methods, where possible, more than one tool should be used [97].

“Professionalism is a complex and dynamic concept that is influenced by a wide range of factors (p185)[112].” Thisthlethwaite and Spencer[112] explain why institutions should invest their efforts in assessing professionalism, particularly given that the concept is ‘complex and dynamic’. By ignoring assessment of professionalism faculty runs the risk of sending the ‘wrong message to the public’ and are “ultimately doing a disservice to society, students and practitioners alike (p186) [112].” The challenge the authors present is the absence of a set of measurement instruments[112], something that this thesis serves to address.

Chapter 2: Methods

This Chapter describes the methods in common for the studies in this thesis, namely qualitative methods. It also explains the rationale behind the chosen methods.

This thesis comes from an interpretive stance, whereby it assumes that reality is socially interpreted and that there are multiple realities or interpretations of an event. This is a constructivist approach, seeking not to find knowledge or fact but rather to construct it.

2.1 Ethics

All studies in this thesis were given ethical approval or were deemed to be exempt from requiring ethical approval by the School of Medicine and Health Ethics Sub-Committee. The ethics status for each study is outlined in the methods section for the Chapter.

2.2 Recruitment

For the studies described in this thesis undergraduate medical students were recruited from Durham University's Phase 1 programme. Students were invited to participate by email, announcements on the Virtual Learning Environment (VLE) and by notices on boards within the department.

Students were provided with information sheets which outlined the study including the following: consent, study design, sponsors, ethical approval, data collection, storage and dissemination.

Students were provided with the opportunity to ask any questions and were aware that withdrawal would not have any negative consequences.

There were no exclusion criteria for participants. This was true for all studies in this thesis.

For studies where participation was compulsory, due to the intervention being part of the teaching, recruitment and consent are outlined in the methods section of the relevant Chapters.

2.3 Qualitative methods

This section described the qualitative methods which have been used in this thesis.

2.3.1 Data collection methods

This thesis utilised four methods for the collection of qualitative data: focus groups, interviews, field notes and free-text comments on evaluation questionnaires.

2.3.2 Data collection

Where qualitative data collection methods were used the following procedures were utilised: all discussions were digitally recorded, recordings were transcribed verbatim by administrative staff, and all participants were given the opportunity to approve transcripts before the data analysis stage.

Interviews and focus groups were semi-structured[153] with the facilitator(s) using a pre-agreed set of interview spines.

Data collection and analysis were in iterative cycles, with each previous discussion informing the next.

2.3.3 Data analysis

All transcripts were analysed using a grounded theory approach[153-158]. Grounded theory is both a methodology and a paradigm. For the purposes of this thesis grounded theory is characterised by the following: iterative cycles of data collection and analysis, constant comparison and purposive sampling.

Grounded theory, developed by Glaser and Strauss (1967), is a means of generating a theory which is ‘grounded’ within or supported by the data[158]. This method uses an inductive approach to analysis, with constant comparison[159].

Before data analysis, I immersed myself in the data by reading all transcripts. This provides a general overview of the data collected. Data were analysed using open, axial and

selective coding[153-158]. Open coding was conducted by researchers independently. At this stage, any data which may have been interest was loosely categorised, or tagged[153]. Throughout this process researchers constantly compared the transcripts. Subsequent data collection may be informed by the codes derived at this stage. Open coding began by the author reading the first transcript, making notes within the margins. Once the transcript had been read from start to finish, the author revisited the transcript. This time, statements of interest, or relevance to the research question were highlighted. For each statement the author assigned a keyword. For the first transcript this was an exhaustive process, to ensure that any information which may be applicable to the study was highlighted. This process was open coding, because, as described by Merriam, “You are being open to anything possible... [160].”

After keywords were assigned to statements, the author reviewed each statement. Keywords were considered in relation to both the statements and the research question and codes were assigned. These codes enabled the author to categorise statements.

After analysing this first transcript I then proceeded to do one of two things; analysis of the next transcript in exactly the same manner, keeping in mind the codes that I had just generated, or scheduled another interview or focus group. The latter was when data were collected in iterative cycles, a characteristic of grounded theory. Where all data were collected simultaneously, grounded theory could still be applied as it is both a methodology and a paradigm. If anything new emerged from the second transcript then the researchers would revisit the first transcript to determine if this theme was present. This was repeated at

all stages of data analysis, and is known as constant comparison[161]. Constant comparison was easily achieved by managing data in NVivo 8.0 for Windows.

After open coding, researchers engaged in axial coding. This was a process of refinement. The relationships between codes were deduced. Here the theory emerged. Axial coding was a process of negotiation between researchers, the same was also true of selective coding. During axial coding any hierarchy between themes and sub-themes was determined and delineated. This is also known as analytical coding, axial coding results from interpretation and reflection on meaning[153].

The final stage was selective coding. During this stage the core themes were decided upon, and illustrative quotes selected. At this point the evidence to support the theory was collected.

Where quotes have been included in the chapters of this thesis, they are not tagged with any identifiable data. This was stipulated in participant consent forms, as students were discussing topics including professionalism, and may be expressing opinions which they did not want to be linked to them.

2.3.4 Reliability

Reliability, with respect to qualitative methodologies, is problematic. This is principally because human behaviour is not static, nor is one opinion more likely to be correct over another. In quantitative terms, reliability is achieved when a study can be repeated yielding the same results[154]. For qualitative research, replication may never be achieved. This does not discredit the data as it should be acknowledged that data can yield many

interpretations. In light of this, the reliability of qualitative data comes from asking whether the results are consistent with the data collected[153].

Consistency in data analysis for this thesis was achieved by using a minimum of two researchers, most frequently three, to analyse data independently.

2.3.5 Validity

Validity is the extent to which findings of a study can be applied to other situations[154], in other words, the generalisability. It is due to this perceived lack of generalisability that qualitative research is sometimes dismissed. Generalising the findings of a qualitative study is difficult, as the sampling is often purposive, and the findings largely contextual. The validity of a qualitative study must therefore come from the transferability of the findings[153]. It has been suggested that the best way to achieve transferability is to therefore paint as full a picture as possible of the context and findings of the study; “a thick description of the sending context so that someone in a potential receiving context may assess the similarity between them and...the study[162].”

Another suggested method for improving transferability is careful selection of the study sample. Maximum variation is a sampling method[153] which will allow for the application of the study to a wider audience. This variation is achievable by increasing the sites used for a study, choosing more participants or selecting a typical sample of participants[153].

In order to make the findings of the studies in this thesis transferable I have, where appropriate, used as large a sample size as possible, used multiple cohorts and included another institution. This enabled the findings to be considered out with the culture of Durham University.

2.3.6 Reflexivity

Reflexivity is the qualitative equivalent of bias[159]. It can be described as the lens through which the researcher looks, or, ‘the process of reflecting critically on the self as researcher, the “human as an instrument’’. Researchers acknowledge their biases, theoretical orientation, dispositions and assumptions regarding the research to be undertaken. Hall and Callery proposed that reflexivity provides “criteria for rigor that make more transparent how data are created within grounded theory studies[163].”

In order to minimise these influential factors, facilitators of focus groups and interviews, who also engaged in data analysis, were chosen because they had no role in the formal teaching and assessment of the subject to be discussed. Facilitators, other than myself, had not engaged with the literature prior to data collection and therefore their assumptions were minimised.

It is important to acknowledge reflexivity to allow the reader to understand how the researcher’s orientation may have influenced the conduct and conclusions of the study.

2.3.7 Rationale

Where qualitative methods were chosen it is because they allow the researcher to understand how people interpret their experiences, what meaning they attribute to their experiences and what context they come from.

Where focus groups have been utilised this is because they allow the researcher to access multiple, diverse opinions in an efficient manner. The information exchange within the group can be interactive and dynamic. This allows exploration of contrary opinions and reflection. If the topic did not require spontaneous exploration of multiple opinions, or the topic was of a personal nature then individual interviews were chosen instead.

Interviews are probably the most widely employed method in qualitative research. Interviews allow researchers to understand participants' personal perspectives and experiences; thus interviews enable a more detailed exploration of the research question.

As with focus groups, interviews have their downside as they rely more heavily on participant memory recall. It has also been suggested that they are more influenced by social context. Unlike focus groups, during an interview the researcher must be engaged. Interviews do not present the opportunity to observe interactions, as the researcher is occupied with questioning.

In some circumstances participants selected to be interviewed instead of attending focus groups. This was primarily due to either timetabling conflicts or the sensitive nature of the topic, as perceived by the participant.

Focus groups provide the opportunity for group interaction, leading to a dynamic exchange of information. This ‘synergistic group effect’ [154] provides information not readily available that individual interviews cannot. The reason for this is that people’s behaviours and attitudes are made within a social context, rather than in isolation[154]. Focus groups may therefore be useful for observing and analysing social influence.

The semi-structured approach to an interview may reduce the influence and role of the facilitator. In addition, this allows the group to take charge of the discussion providing an opportunity for the facilitator to observe the interactions between participants.

As with all methods, focus groups have their limitations. Focus group responses can be more extreme than those collected in a survey[154]. During focus groups the dynamic can be altered by one individual who may be influential or opinionated. In some situations, if an extreme attitude is held by an outspoken person, the focus group dynamic can shift towards extremism. The outspoken participants may silence any contrary views; Noelle-Neumann (1984) described this as the Spiral of Silence Theory whereby people shape their opinions to prevailing attitudes about what is acceptable[164]. The Spiral of Silence Theory was originally described in relation to the media. However, it is equally applicable in the focus group context. Participants may be unwilling to publicly express their opinion if they believe

themselves to be in the minority. Thus, the participants become more marginalised, resulting in them speaking less and therefore they spiral into a fully marginal position. The same participants may be more vocal if they believe they are a part of the majority.

Focus groups may not provide as much depth nor breadth as interviews. Due to multiple opinions and time constraints, the number of questions that can be covered may be limited. Focus groups generate major themes, however, are often not as suitable for identifying subtle differences in opinions [154].

Field notes were utilised as they provided a written account of an observation (Merriam).These notes were written as soon after the observation as possible in order to ensure accurate information recall.

Where appropriate field notes included the following, as suggested by Merriam (2009)[153]:

- Descriptions of the setting, participants and activity
- Quotations or the substance of what people were saying
- Observer comments (anything of interest)

These first hand accounts of events are powerful when combined with other qualitative data, for example focus group transcripts or evaluation data. The impact of having an observer present is discussed with the relevant Chapters; however, it is worth noting at this point that the presence of an observer has an effect on any situation. This is true whether the observations are conducted as a participant or an observer. Any interaction leads to

some degree of distortion of the situation[153]. To reduce the level of distortion from my presence as a participant observer, I refrained from making notes during any session. This enabled me to better socialise into the setting without being consciously badged as a researcher by the participating staff and students.

Evaluation questionnaires provided free-text data which were collated and subsequently analysed. As with focus group, interview and observation transcripts, evaluation questionnaire data were analysed using grounded theory.

Free-text responses do not provide as rich a data source as other sources of qualitative data, such as focus group transcripts. Free-text responses, such as those obtained from evaluation questionnaires in this thesis, provide rapid responses when time constraints exist. I acknowledge that these responses may represent knee-jerk reactions to the question being asked and often lack context. However, they can be a valuable source of information from participants who may not have time or be willing to contribute to interviews of focus groups. Free-text responses are particularly useful for justifying quantitative answers on questionnaires or for canvassing participants' thoughts on interventions. I have found free-text options on evaluation questionnaires provides opportunistic data from intervention participants who may otherwise may not contribute to qualitative aspects of my research.

All qualitative research in this thesis utilised a semi-structured approach to data collection. Semi-structured interviews and focus groups allowed flexibility during questioning.

Participants were able to lead the discussion, framed around their understanding of the issues and events. Additionally, this approach enabled the researcher to explore unexpected issues as they arose.

2.4 Quantitative methods

Where quantitative methods have been used, they are described in the relevant chapters.

Chapter 3: Virtual Human Dissector as a Learning Tool for Studying Cross-sectional Anatomy

3.1 Background

3.1.1 Contributions to this Chapter

This study has been published in Medical Teacher [69] and presented at the Association for Medical Education in Europe Conference in Norway, August-September 2007[165].

I collected data for the study, contributed to the data analysis, and wrote the manuscript which was published in Medical Teacher[69]. Dr Leo Donnelly collected data, contributed to the data analysis, and to revisions of the manuscript. Dr Patten and Mrs Pam White contributed to the data collection and to revisions of the manuscript. Professor John McLachlan advised on study design.

3.1.2 Conceptual Framework

This study was a comparison of computer-based and traditional anatomy teaching methods of medical undergraduates and their consequent ability to interpret two-dimensional images similar to the computer tomography (CT) and Magnetic Resonance Images (MRI) they will encounter clinically. These scans rely on digital technologies which produce images showing constrast between different tissues. CT and MRI images are traditionally presented as cross-sectional views of the subject's anatomy.

The Virtual Human Dissector, at the time of this study, was a new software package. It was being used in anatomy teaching at Durham University for the purposes of projecting images of anatomical structures onto a student volunteer. The projected viscera could be scaled to the volunteer and then palpated by one of their peers. Students were enthusiastic about the use of VHD. This resulted in Dr Patten implementing further VHD learning activities into the curriculum. These activities were both instructor and student-led. Staff demonstrated cross-sectional anatomy using VHD or students completed questions based upon its use. The aim of these learning activities was to encourage and engage students in cross-sectional anatomy and to give them greater familiarity with both the anatomy and using computer software.

This study was devised to assess whether there was a difference between the effectiveness of the VHD and traditional anatomy teaching resources, such as prosections and plastic models, for students studying cross-sectional anatomy.

3.1.3 Study aims

The aim of this study was to ascertain whether or not VHD is effective as a SDL tool to facilitate students' ability to interpret cross-sectional images and their understanding of the relationships between anatomical structures.

3.2 Introduction

Within modern day healthcare settings there is increased use of technologies such as CT and MRI scans. The Canadian Institute for Health Information reported that the supply of MRI and CT scans and the number of scans performed have risen dramatically in the past decade; a trend mirrored worldwide (CIHI, 2006). In England during 2006-7 there were 2,728,119 CT and 1,257,972 MRI scans performed (158% and 219% increases, respectively, over the previous 10 years; Figure 3 - Department of Health Hospital Activity Statistics).

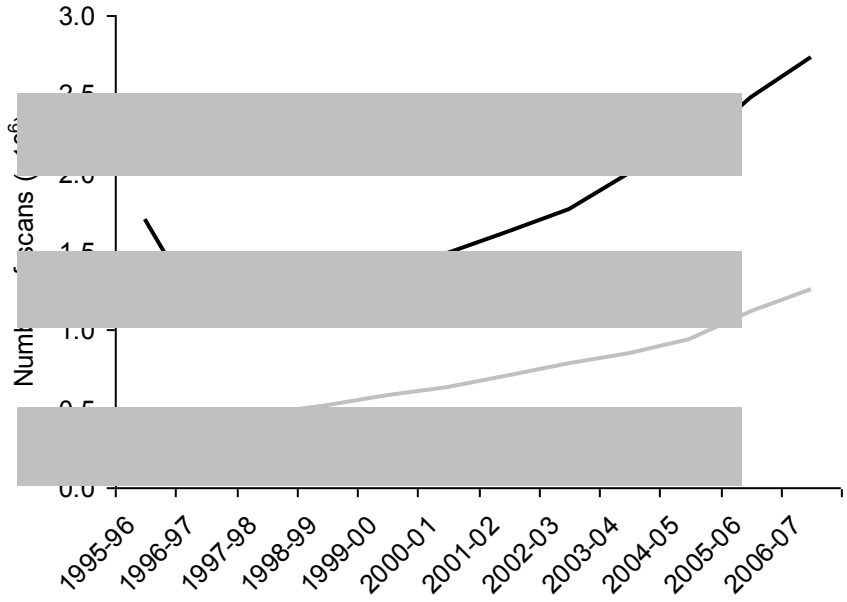

Figure 3: Numbers of CT and MRI scans carried out in English hospitals since 1995-1996. ***Source: DoH Hospital Activity Statistics; Form KH12***

Dawes et al. (2004) [166] have reported a marked increase in successful interpretation of radiographs by medical students following training, which may be extrapolated to cross-sectional images. In light of this, the curricula of Medical Schools may have to increasingly focus upon cross-sectional anatomy, as well as to conform to the guidelines from the GMC's *Tomorrow's Doctors* which urges medical schools to use new technologies in teaching[20]. There are many software packages available to assist in the teaching of cross-sectional anatomy, such as Acland's Cross-sectional Navigator, NCI Flicker (which compares images from different internet sources), the Segmented Inner Organs Data of the Visible Human Male, and others reviewed by Jastrow & Vollrath (2003)[67].

Another such programme is VHD, which is an interactive anatomy programme based on the National Library of Medicine's Visible Human Project[64]. It was developed at the University of Colourado, which acquired the specimens and images that comprise the Visible Human DataSets™.

The Visible Human Project is a complete, anatomically detailed, three-dimensional representation of an entire male and a partial female human body. CT and MRI scans were taken before thin (1mm or less) serial transverse cryoslices were taken from each gelatin-embedded cadaver (1871 slices from the male, 5189 from the female), and digital and analogue photographs taken of the remaining part of each body. The resulting digital image library allowed for (laborious) identification and multilingual labelling of every visible structure in the data set[64].

As described in the conceptual framework, within undergraduate medical teaching at the University of Durham, this software is used for both formal teaching sessions and the students' self-directed study. Formal teaching sessions include use of VHD with a portable data projector to superimpose anatomical images onto volunteers[167]. These images can be scaled to the volunteer, using their bony landmarks as a guide, and the projected structures can be subsequently palpated by peers.

Prior to this study VHD had not been used by the students taking part in the study. Other student groups at Durham use VHD for SDL, allowing them to study cross-sectional anatomy independently or as a learning activity within an anatomy practical. Drake (1998)[5] suggests that self-directed, lifelong learning should be emphasized within a medical curriculum and in agreement with this, SDL constitutes part of most anatomy practicals at Durham where its purpose is to provide students with a timetabled session in which they may tackle learning outcomes alone, rather than being taught didactically. During SDL the student takes the initiative and responsibility to complete the prescribed learning activity and a debriefing ensures that the students take away the correct information.

3.3 Methodology

3.3.1 Ethics

This study was declared as exempt from requiring ethical approval by the Chair of the School of Medicine and Health Ethics Sub-Committee at Durham University, Professor Martyn Evans.

3.3.2 Recruitment

Participation in this study was compulsory as the intervention was part of a scheduled anatomy class which all students must attend. Testing often takes place during teaching sessions, therefore completion of the tests described in this study was also compulsory.

3.3.3 Study Design

This study employed a cross-over design, depicted in figure 4.

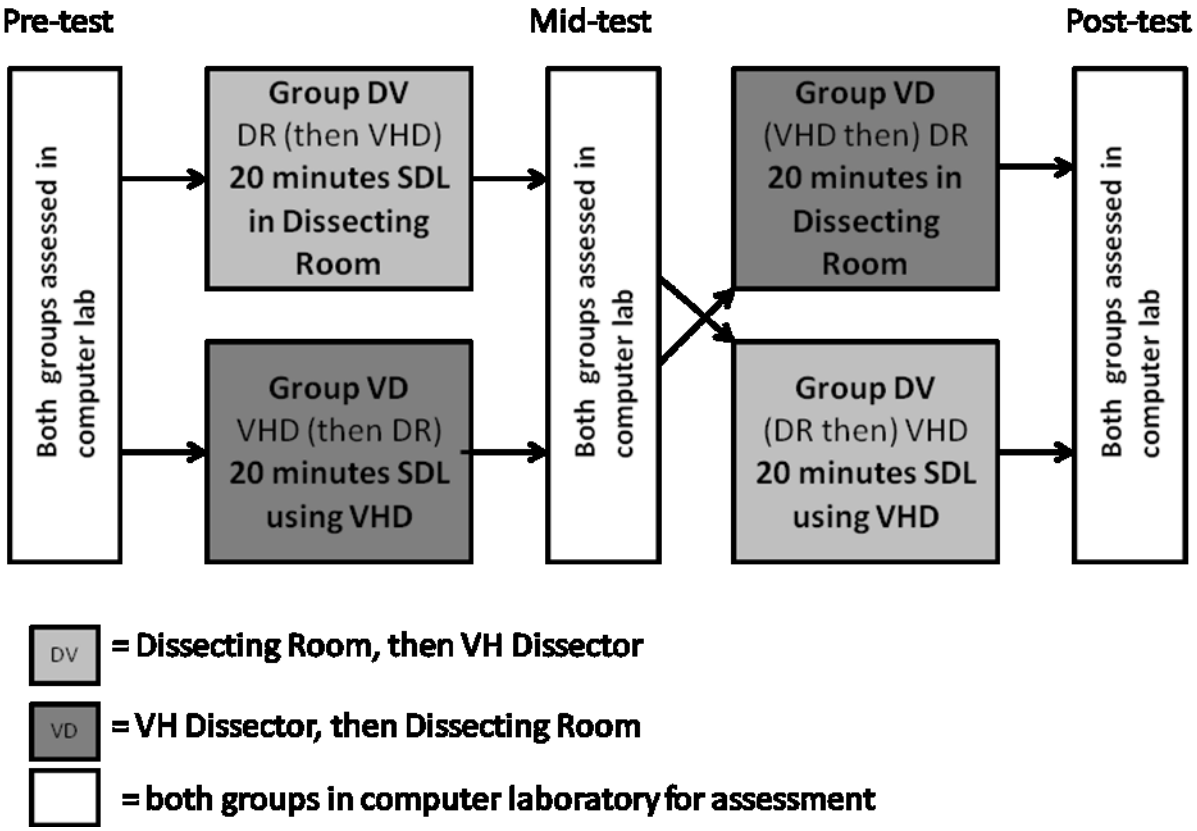

Figure 4: The cross-over design of the study.

Data collection took place during the academic year 2006-7. Eighty-nine first year medical students participated. The cohort was divided into two groups (groups 1 &2); group membership is randomly assigned at the beginning of each academic year.

Students were assessed on their knowledge of cross-sectional abdominal anatomy prior to the self-directed learning session. This formative assessment was part of an Anatomy

practical within the NME (Nutrition, Metabolism and Endocrinology) strand. The pre-test served to determine that the random distribution of students between groups was equal with respect to prior knowledge of cross-sectional abdominal anatomy.

The format and method of analysis of the pre-test was also utilised for the mid and post-tests. Students were required to answer 12 computer-based multiple choice questions (MCQs) identifying numbered anatomical structures on 3 photographs of body slices on (the pre-session test). The photographs were of cadaveric specimens, but were not the images used in the VHD software. The test was constructed using Microsoft Powerpoint 2003 and delivered using Blackboard™ on the University's Virtual Learning Environment (VLE). The Virtual Learning Environment (VLE), as used by Durham University, is an integrated set of online tools, managed resources and databases, existing as a coherent system[168]. Data were captured automatically and were subsequently downloaded to Microsoft Excel 2003. Scores were computed as '0' for an incorrect answer, and '1' for a correct answer. For each student a mean test score was calculated (total divided by 12), ranging between 0 and 1. Test data were anonymised.

This study was part of an abdominal anatomy practical session. The session had 4 stations; 2 small group teaching stations using prosections and 2 self-directed study (SDL) stations. The self-directed study stations were the subject of this study. One SDL station was based within the DR, where students had access to textbooks, prosections and plastic models. The other was based in a computer lab where students had access to the Virtual Human Dissector software and were able to scroll through and study cross-sectional images of the abdomen.

Group 1 (n=44; designated VD) remained in the computer suite and were allowed 20 minutes of self-directed study on abdominal cross-sectional anatomy using VHD. Students had not been given the opportunity to use VHD prior to this session, but had been given a short introductory seminar on its use. The only instruction given to the students was that a mouse-over function could be used to identify the vertebra, and therefore the vertebral level, of a cross-sectional image.

Group 2 (n=45; designated DV) were escorted to the neighbouring DR and simultaneously undertook a more traditional (and familiar) 20 minute self-directed learning (SDL) session using prosected cadavers, models and textbooks. Students did not receive any teaching during this session. Students have familiarity with this mode of self-directed study as they have up to 2 hours SDL timetabled per week in which they can access the DR and its contents.

Following this first session of SDL all students returned to the computer suite and completed a second computer-based test (the 'mid-session test'), using the same MCQs as the pre-test, but reordered.

The groups were then 'crossed-over'. Group 1 was removed to the DR for 20 minutes SDL whilst group 2 remained in the computer suite for SDL using VHD.

Following this second session of SDL all students returned to the computer suite and completed a third and final computer-based test (the 'post-session test') using the same MCQs, again reordered.

Test scores were analysed using anonymous Student ID numbers. Anderson-Darling tests indicated that each set of results was normally distributed. Comparisons were made with two-tailed Student's t-tests using Microsoft Excel 2003, which generated p-values reported in tables 2 and 3.

3.4 Results

Both groups' test performance increased between the pre- and mid-session test and between the mid- and post-session test. The differences in mean score between the two groups for each test were not significant at any stage.

		n	Mean score	SD	p-value	Sig.
Pre-session test	DV group	44	0.398	0.213	0.797	ns
	VD group	45	0.409	0.209		
Mid-session test	DV group	44	0.559	0.139	0.918	ns
	VD group	45	0.556	0.150		
Post-session test	DV group	39*	0.641	0.108	0.857	ns
	VD group	39*	0.637	0.100		

Table 2: Comparison of pre-, mid- and post-session test results for DV and VD groups

**5 students from the DV group and 6 students from the VD group were given permission to leave early (before the post-session test), in order to attend a community placement. Excluding their results entirely from the statistical analysis does not change the probability values.*

No significant difference occurred between the two groups at any testing stage.

		n	Mean score	SD	p-value	Sig.
DV group	Pre-session test	44	0.398	0.213	<0.001	***
	Mid-session test	44	0.559	0.139		
	Mid-session test	44	0.559	0.139	<0.001	***
	Post-session test	39	0.641	0.108		
VD group	Pre- session test	45	0.409	0.209	<0.001	***
	Mid- session test	45	0.556	0.150		
	Mid- session test	45	0.556	0.150	<0.01	**
	Post- session test	39	0.637	0.100		

Table 3: Comparison of test results within each group

Both groups improved their test results between the pre- and mid-session and the mid- and post-session tests. There was, however, no statistically significant difference between the two groups at any stage.

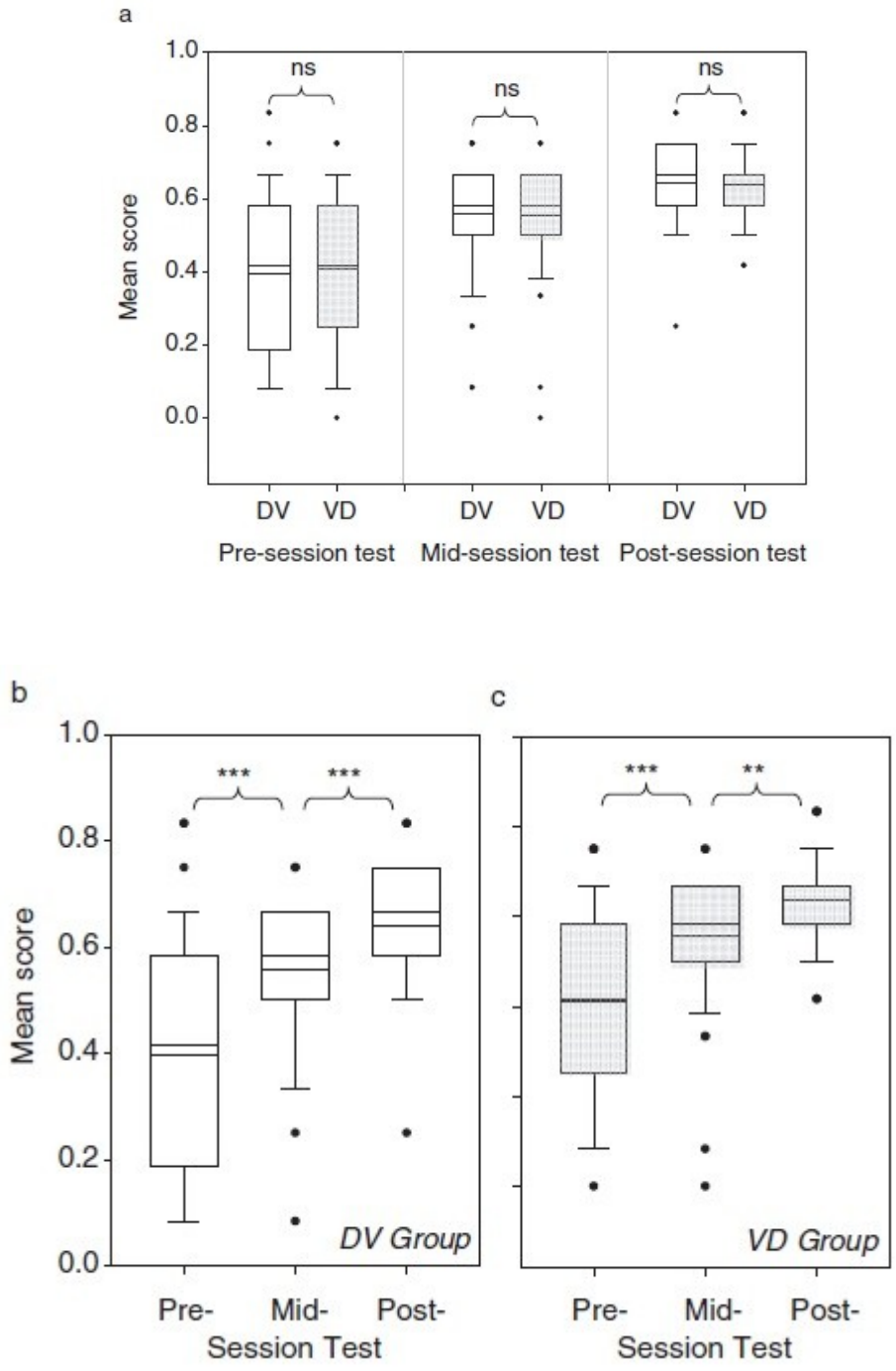

Figure 5: Box graphs indicating distributions of scores. Boxes indicate 25th and 75th centiles, and 'error bars' the 10th and 90th centiles. The fine horizontal line is the median, the bold line the mean. Dots indicate outlier values.

a. Comparison of group scores within each test

b. Comparison of DV group scores over all tests

c. Comparison of VD group scores over all tests

(ns = non-significant; ** = p<0.01; *** = p<0.001)

3.5 Discussion

As there was no significant difference between the two groups at any particular testing stage it is suggested that VHD provides an alternative rather than an improved learning experience. Also, the timing of this learning experience (i.e. whether it preceded or followed the DR session) did not affect the learning outcome (test score).

As the test results increased equally for both groups between the mid- and post-tests, i.e. the results were additive rather than reaching a plateau, it is possible that learning has taken place by different pathways. Such pathways may include surface or deep learning, kinaesthetic, visual, auditory or tactile learning. This possibility needs to be explored further.

The results show that the VHD is as effective a learning tool for self-directed learning as prosections. Durham finds VHD a useful adjunct to formal anatomy teaching with cadavers. 'Users can examine the anatomy of individual viscera, structures or entire organ systems simultaneously in both 3D and 2D views' [167].

VHD allows some structures that are often difficult to see within the cadaver to be more easily visualised, e.g. neurovascular structures within an organ can be followed by scrolling through serial sections. VHD also allows students to follow the changing profile of structures at different cross-sectional levels. An example given by Patten (2007) is the

transition of the descending aorta to the common iliac and then external and internal iliac arteries[167]. VHD can also show the 3D relationship between structures.

As mentioned previously, the marked rise in the use of CT and MRI has increased the need for understanding of cross-sectional anatomy. Such an understanding cannot be produced or indeed learned on a purely dissection or prosection-based course [40]. VHD can be an excellent, interactive way to introduce cross-sectional anatomy without the use of cadavers.

I do not propose VHD as a substitute for formal teaching, although I do suggest it can be an invaluable resource for students who experience difficulties with cadavers or the DR environment. VHD can be readily accessed on a number of university computers, not only enabling students to access it in their own time but also providing an alternative study environment for students who are anxious about studying in the dissecting room. Skidmore (1995) highlighted that a fear of death may be oppressive for students, and may be correlated to poor academic performance[40]. VHD provides students with reservations toward the DR with a learning opportunity in an alternative environment which may alleviate any fears and help improve performance.

VHD may also provide an alternative learning resource for students when the DR is not available due to teaching constraints or unavailability of staff, a problem that is expected to increase in years to come with declining numbers of graduate students going on to teach anatomy[41]. The advantage of computer assisted learning is that it allows students access to course material off-site. Students are potentially able to study cross-sectional anatomy at

any vertebral level, at any time of day. It is not practicable to have cross-sections of specimens available to students at any desired vertebral level within most dissecting rooms. However, packages such as the VHD enable this. VHD is also a practical alternative for students who find death oppressive and thus struggle, and do not find the DR an environment conducive to successful study.

Encouraging self-directed learning is, in my opinion, important as students need to acquire the skills necessary to teach themselves and become life-long learners. The problem with the traditional teacher-centred scenario common to anatomy teaching is that students fail to take responsibility for their own learning which may result in a loss of motivation[169]. Offering alternative approaches to SDL, such as the option to use the VHD, may enable institutions to engage students more readily with SDL.

The use of the VHD enables institutions to adhere to guidelines set out by the GMC's report on *Tomorrow's Doctors* (2003)[20] .Such guidelines include the following:

"Medical schools should take advantage of new technologies to deliver teaching" and that "clinical and basic sciences should be taught in an integrated way throughout the curriculum."

In this respect the VHD has been found to be useful as it allows anatomy to be taught simultaneously with clinical skills, such as percussion and auscultation, when projected onto student volunteers or life models [167].

Advantages VHD has over cadaveric dissection include that body systems can be viewed in 3D in a variety of orientations and a virtual slice can be reviewed any number of times, with structures added or removed. Students can use VHD in 2D to view transverse sections, and neurovascular structures in particular can be followed by scrolling through sections. This scrolling action produces imagery which appears as if it were animated. Ruiz and colleagues described how computer animations, utilised in medical education, can depict dynamic changes over time and location[170]. Such animations can illustrate phenomena and concepts that might otherwise be difficult for students to visualise. However, they warned educators that animations may not always be effective and that an understanding of the principles that govern their use is required. It must also be acknowledged that the benefits of animations may also vary according to individual learner characteristics such as prior knowledge and spatial ability; a point worth noting with respect to the use of the Virtual Human Dissector and other computer software which may not necessarily involve an animation.

Learning to interpret images is a troublesome process for medical students. Thus, including modalities which introduce cross-sectional anatomy (such as the VHD) into the pre-clinical curriculum enables students to familiarise themselves with this difficult concept early on in their careers. Furthermore, the clinical relevance of cross-sectional anatomy should be sign-posted to students and its incorporation into the curriculum in this context will cause

students to weight the associated anatomy more highly, as they understand its relevance to the practice environment.

The use of knowledge tests immediately after SDL may be advantageous to students; Fry et al (2008), state that learning of material is consolidated if students are given the chance to use the information within a short time of its presentation[171]. Furthermore, assessment is a driver for learning. This was described by Marton (1984), Ramsden (2003), and specifically with respect to medical students by Pandey and Zimitat (2007)[172, 173]. Assessment is extremely important in vocational courses such as medicine since assessment continues until a practitioner retires.

Post-teaching testing is a valuable way to improve retention of knowledge [174-176] as it has been shown to slow decay of knowledge [175]. Karpicke and Roediger found that it is the *"repeated retrieval induced through testing, and not repeated encoding during additional study, that produced large positive effects on long-term retention"[175].* Cognitive psychologists have named the phenomenon as the testing effect. The testing effect shows that testing directly affects learning by promoting better retention of knowledge [176]. Bloom & Hastings (1971) said that, *"Formative assessment could become a powerful tool shaping students" performance"[177].* The widespread use of VLEs is driven by the *'ubiquity of computer-based activities in education'* and growing pressures for increasing *'quality and quantity of educational efficacy and student support'[168]*. Delivery of formative assessments on the VLE, as used in this study, may contribute to this.

"Within less than two student generations, communications and information technology has been repositioned as an integral component of the medical school environment" [168]. Current students are the technology generation. Delivering teaching using modalities with which they are comfortable such as computer software, could potentially lead to better acceptability from students. Acceptability is a consideration when implementing any change into the curriculum.

It could be argued that the VHD and similar products are related to simulators. For example, VHD enables the students to track the changing profile of the anatomy and, to a limited degree, manipulate the angle at which it is viewed. This is similar to some simulation packages. Simulation allows the novice and the expert to practise and develop skills with the knowledge that mistakes carry no penalties [178], and may aid subsequent recall, as will be described in Chapter 5. Although the VHD is not always regarded as a simulator, which for example are usually thought of as high-fidelity simulated patients, it does enable students to master interpretation of cross-sectional anatomy and related medical imaging in a safe environment, absent from the pressures of the clinical context. In 2007 Jason stated that, *"Your medical students will most likely become clinicians. Unless you can help them see the connections between your offerings and their future work, you may not capture nor sustain their interest"[179].* Teaching anatomy in a clinical context is imperative to this, it enables the anatomy taught at an undergraduate level to be connected to students' future work. The VHD is a mechanism by which the anatomy taught can be given clinical context. The VHD could be used as adjunct to other imaging techniques such as MRIs and CTs when teaching anatomy.

There are, as with all modalities, disadvantages associated with using computer packages. One risk is that students will no longer access the dissecting room. This could be problematic for a number of reasons, for example they may become incapable of identifying structures in a whole cadaver from a bird's eye view, attendance may drop and communication skills diminish due to a lack of contact with tutors and peers. Thus, computer packages are best incorporated into a blended curriculum.

Their implementation into a blended curriculum would require prior knowledge of the students and their approach to learning, explict course learning objectives and an information technology infrastructure. Institutional access would be required for students without remote IT access, plus IT support for all users. Guidance on best practice, and what level of study or usage would be beneficial to students.

The take-home messages from this study are that:

- VHD may be a useful adjunct to formal anatomy teaching as students can access it in their own time
- VHD may be as effective as using prosections for self-directed study
- VHD can be utilised as a practical alternative to anatomy teaching for students who encounter difficulties with cadavers and / or the dissecting room
- VHD is valuable in aiding students to interpret 2D and 3D diagnostic images
- VHD appears to have high acceptability with students

3.5.1 Limitations

This study was only conducted using one cohort on one occasion, thus generalisability and transferability may be questionable. Long term retention of knowledge using the VHD was not addressed. Results may have been different if students had as great a familiarity with the VHD as they did with the typical self-direction learning session.

3.5.2 Future work

Future work may include tracking the number of hours students spend accessing the VHD and correlating this with exam performance on questions which require interpretation of cross-sectional anatomy.

It may be of interest to ascertain whether there is any difference in the VHD's effectiveness for different body regions, for example the head and neck which have a more complex structure, making cross-sections taken at this point more difficult to interpret. Using VHD within this region may be more advantageous than in other regions where the anatomy is less dense.

A more relevant test of a student's knowledge in any future work might be to use CT and MRI scans rather than images of cadaveric cross-sections. A second post-session test one month, say, after the self-directed session, might determine retention time between the two groups.

It might also have been advantageous to incorporate a focus group post-test, using student volunteers. This would enable investigation of the preferences of the students towards either of the SDL methods and explore the idea of usefulness of VHD for students experiencing difficulties with cadavers. Ogilvie *et al* (1999) have reported that medical students like the use of computer administered exams, and such exams may accentuate the learning experience[180]. This could be investigated during the focus group.

Finally, it would be beneficial to ascertain whether or not use of the VHD has any long term effects on recall. In any future repetitions of the study, an additional post-test would reveal any differences between study using prosections and the VHD on retention of knowledge.

Chapter 4: Body painting as a tool for learning anatomy in an undergraduate anatomy curriculum.

4.1 Background

Part one of this Chapter has been published in the journal *Anatomical Sciences Education.* Elements of this Chapter have been presented at numerous conferences including the American Association of Anatomists, the Anatomical Society of Great Britain and Ireland and Create[85, 181-183].

4.1.1 Contributions to this Chapter

I was responsible for the design, data collection and analysis for both parts of this study.

Professor John McLachlan was co-author on the publication[90] of part one of this study in the journal *Anatomical Sciences Education*. He contributed to data analysis and editing of the final manuscript.

Dr Debra Patten contributed to the study design for part two of this study.

Dr Debra Patten, Dr Michael Griksaitis and Mrs Pamela White helped facilitate the teaching sessions utilised in the interventions described in this study.

4.2 Introduction

Body painting within Medical Education is described as painting internal structures on the surface of the body with a high degree of detail [88].This contrasts with simple line drawings as used in conventional practice, which can be dated back for at least a century[184]. Body painting proves particularly useful as a method for introducing surface anatomy in medical teaching, and complements the teaching of clinical skills such as palpation and auscultation[88-90]. Furthermore, it is thought that the active and kinaesthetic nature of body painting, coupled with the strong and highly memorable visual images of underlying anatomy, contribute to its success as a learning tool [39, 88-90]. Body painting has also been reported to be a valuable tool for diminishing the apprehension often exhibited by students when conducting peer physical examinations [89]. Similarly, Skidmore (1995) suggested that a fear of death may be oppressive for students studying in the dissecting room and may be correlated to poor academic performance; the use of alternative teaching methods, such as body painting, may therefore be beneficial to students who struggle with cadaveric work[40].

Students have reported body painting as a highly motivating exercise. Its use in large group teaching is feasible. In my experience its main advantage appears to be the creation of what could be called ‘learning landmarks’: vivid experiences which are memorable in themselves, and which then provide access to the educational content associated within that context [90]. In particular, students acquire a good understanding of dimensions and positions of anatomical structures using this method whilst studying [88]. Literature is in accordance with the idea that body painting is a highly memorable experience, which gives students an

appreciation of the links between the visual, tactile and auditory aspects of human anatomy [36].

As described in Chapter 1, since the publication of the *Tomorrow's Doctors[20]* (GMC, 2003), there has been a national change in the medical curricular structure within the United Kingdom (UK). Two statements from Tomorrow's Doctors[20] (GMC, 2003) summarise the recent changes in the UK:

- "Modern educational theory and research must influence teaching and learning. Medical schools should take advantage of new technologies to deliver teaching."
- "The clinical and basic sciences should be taught in an integrated way throughout the curriculum."

Raftery (2007) suggests that the recommendations in *Tomorrow's Doctors* have been detrimental to anatomy teaching and have resulted in a lack of medically qualified teachers of basic sciences, increased workload demands and the increased commitment to postgraduate surgical training required [185].In the period following the publication of Tomorrow's Doctors[20] there has been a loss of gross anatomy teaching time [22], and a reduction in the detail taught[186]. There has been greater integration between anatomical disciplines and clinical skills. Body painting is one means by which Durham University has further integrated gross anatomy and clinical skills. Barrows et al. (1968) remarked that medical students were often not given the chance to relate the anatomy that they had learnt within the laboratory directly onto a living body. The authors called for students to be able to palpate structures and see the structures studied in the cadaver move and function. "In this manner a solid learning bridge is formed between anatomical knowledge and further pursuits in physical diagnosis and clinical medicine is formed[74]."

Historically, anatomy has been taught by dissection. Cadavers are useful for studying the anatomy of large organs[24]and give an overview of spatial orientation[7, 119]. However, anatomy is most commonly encountered by medical practitioners in the form of living anatomy and medical imaging[39]. Students should also be encouraged to engage in teaching activities with an emphasis upon living anatomy and body painting is one such activity. Furthermore, “clinicians often blame anatomists for teaching students too many details and not enough clinically relevant structures” [18], body painting allows the clinical relevance of the gross anatomy to be emphasised to students. Surface anatomy is a way of bringing cadaveric anatomy to life[76] and body painting falls into this category.

Despite this, body painting still appears under-researched. Part one of this study, using qualitative methods, describes the perceptions of undergraduate medical students towards body painting, and in particular their perceptions of its efficacy as a learning tool. Part two of this study measures the impact of body painting as a teaching tool, using a cross-over design.

Before the attitudes of the students towards this pedagogical approach can be considered, the context in which it is used ought to be outlined. Durham utilises body painting in both gross anatomy and clinical skills teaching, which are often taught concurrently. Students paint viscera, bony landmarks and neurovascular structures amongst others. Specific examples are the muscles of facial expression and the terminal branches of the facial nerve, regions of abdominal referred pain, and bones of the hand. Painting is compulsory, as it forms part of the surface anatomy curricula; however being painted upon is optional.

Body painting sessions are usually part of timetabled anatomy and clinical skills teaching. For any one student an anatomy practical usually comprises 4 thirty minute stations. Each station usually addresses one aspect of the body region being taught, for example, a teaching session on the anterior thigh may have the following stations: musculature, nerve supply, blood supply and osteology. When body painting is introduced into a practical it may form one of the 30 minutes stations, or become a double station if a clinical skill is being taught simultaneously. The stations for the anterior thigh session may then become; musculature, neurovasculature, osteology and a body painting station which teaches the distribution of dermatomes.

4.2.1 Conceptual framework

Before commencing my doctoral studies I was employed at Durham University as an anatomy aide. This meant I was responsible for the technical aspects associated with the functioning of an anatomy department. One of my responsibilities was to set-up and facilitate body painting sessions. Whilst working as the anatomy aide I observed at first hand the students' enthusiasm for body painting.

My anatomy education was dissection based, so body painting was a novel experience. As a relatively new practice, the literature on body painting was extremely limited[88, 89]. I therefore decided to explore students' perceptions of body painting. This was due in part to my own curiosity, but also due to questions I had been asked by colleagues at other institutions. Colleagues were concerned about students' willingness to engage with an activity that involved levels of undress, as well as the reactions of students to an activity which could be considered puerile. The aim of this study was to ascertain why the students were so

enthusiastic about body painting, but also to determine what factors needed to be given attention by institutions considering piloting body painting.

Part 1. Students' perceptions of body painting as a tool for learning anatomy.

4.3 Methodology

4.3.1 Ethical Permission

Ethical permission for this study was confirmed by the Ethics Committee of the School of Medicine and Health at Durham University.

4.3.2 Context

Durham University is a Phase One undergraduate medical programme, that is, it educates first and second-year students only, from a five-year curriculum.

4.3.3 Recruitment

Over a two year period (2007-2009), all first and second year medical students from the Phase One programme (305 students) at Durham University were invited to participate in a qualitative research project. Invitations were sent via email, via the Virtual Learning Environment (VLE) and posted on notice boards. Invitations outlined the project, and included

a copy of the consent form and a participant information sheet for study. All students were invited to e-mail or arrange a face-to-face discussion with the principal researcher (GF) to address any queries or concerns regarding the process.

4.3.4 Consent

Information sheets outlining the purpose of the project, data collection and its subsequent protection and how any findings would be disseminated were issued to all students which. Individual consent forms were signed by all participating students. Students were reassured that non-participation or withdrawal would not incur any negative consequences.

4.3.5 Data collection

Data were collected over the two year period from 2007 to 2009 at Durham University, UK. A total of 133 students participated; these students were from 3 cohorts. The demographics of participants are shown in tables 4 and 5. First-year students in 2007 were the same cohort as the second-year students in 2008, however, these were not the same individuals, and therefore there were 133 unique participants.

	Year 1		Year 2	
	Males	Females	Males	Females
2007-8	41	23	7	4
2008-9	17	21	11	9

Table 4: Focus group participant demographics

Total Males	76
Total Females	57
Total 2007-8	75
Total 2008-9	58
Total participants	133

Table 5:Total numbers of students by sex and cohort.

The author (GF) facilitated all semi-structured focus groups, of which there were 24 in total. Groups were mixed with regard to gender but not year of study. Focus groups were the chosen method of data collection as they allowed authors to explore multiple opinions simultaneously. Each focus group had, on average, 6 participants.

I felt that thematic saturation was reached, as after 20 focus groups no new themes emerged. Although 24 focus groups may seem excessive, a further 4 focus groups were conducted in case any divergent opintions were to be voiced. This was because the themes emerging appeared to be universally positive and I was aware that not all students were as supportive of body painting. I was keen to explore the opinions of these students.

The discussions were digitally recorded and transcribed *verbatim*. Transcriptions were carried out by administrators within the department. Transcripts did not identify participants' details in any way. In order to confirm the participants' interpretation of the data, the author (GF)

engaged in member checking [159] both during the focus groups and data analysis. All students were given the opportunity to approve transcripts before publication.

Data were collected and analysed in iterative cycles. Data were analysed using a grounded theory approach [155, 156, 158]. Both the author (GF), and her PhD supervisor (JMcL), analysed the data. Grounded theory was chosen for data analysis as it allowed an understanding of the students' attitudes towards body painting to be generated which was 'grounded' within the data. To ensure reliability, open coding was carried out by the authors independently. Open coding was inductive, in other words codes were not pre-set, but emerged from reading the data. Axial and selective coding of the themes was conducted by the author (GF) and her supervisor (JMcL) through negotiation[153, 157]. Throughout data analysis there was constant comparison[155]. Constant comparison ensured that any codes identified were consistent across the data set. This method served to determine 'similarities and differences' by comparing one segment of data with another[153], in this case all other transcripts. Similar data were then assigned to the relevant code (theme).

4.3.6 Reflexivity

Before considering the results it is important to mention reflexivity, the influence a researcher brings to a project [159]. For the purpose of discussing reflexivity, this section will now change from the third person to the first person to enable a discussion of how as an individual and a researcher I feel I may have influenced the study.

As with all qualitative research, it is important to acknowledge my background as a researcher within the specific context of this Chapter. This background will contribute bias of which I may or may not be aware, but equally, my background enables me to engage with the research.

For part one of this study, I was the researcher responsible for facilitating focus groups. Since I have studied anatomy/am a trained anatomist, the students were able to discuss body painting and anatomy using scientific terminology, a language comfortable for both the facilitator and participants. However, since I also had no formal role in the assessment of anatomy for any of the participants, this reduced the effect of any perceived power relationships which may have affected the data collection. Subject knowledge (anatomy) and practical experience of body painting benefited me as a researcher as I understood the students' experiences, but it also meant I had presuppositions with regards to the data collected. For example, from observing teaching sessions I had already noted the impact of body painting in colour on the students. To avoid this biasing results, pre-decided interview spines were utilised. This allowed students to discuss colour, should the issue arise spontaneously, without the facilitator influencing any responses from participants.

Part 2 of this study describes a participant observation. My background as an anatomist allowed me to socialise into the teaching session. This makes my presence legitimate in the eyes of the participating students. Furthermore, my anatomy and medical education background provided insight and expertise which inform my interpretation and perceived understanding of the research question and findings. Particularly, this enabled my understanding of the pedagogical implications of the findings, informing my future practice.

As an anatomist, students engaged with me within a teaching session. They recognised my content knowledge. As I was not the curriculum lead and did not contribute to the students' assessments I found that this enabled me to participate in more candid discussions with the students regarding their perceptions of the session content and delivery.

At this point in my research I felt that students are also aware of my track record in publishing research. They have witnessed interventions, such as the use of scrubs in teaching [187], and seen publication of results and implementation of successful interventions within the curriculum. I feel that this increases students' willingness to both engage in the research project, but from a reflexive stance, increases their willingness to engage with me as a researcher. This is either by discussion during participant observations, by completing evaluation materials, or by attending focus groups.

As a PhD student involved with teaching on a formal basis, students are aware of my presence within the department. All participants have been taught formally by me. This has potential to limit their interactions, and the transparency of such exchanges, with me during an observation. However, I feel that as students are aware that I currently have no formal role in assessment and their academic progression, their comments remained forthright.

My background, as previously alluded to, does not only impact upon how participants engage with me but also how I interpreted data collected. As this study employed content analysis using grounded theory, my choice of method acknowledged that as a researcher I have

presuppositions of which I am not aware. Grounded theory is an inductive process by which I am seeking to gather data to build a theory, rather than deductively testing a hypothesis, thus potentially reducing the bias to which data may be subject.

4.4 Results

Five principal themes emerged from the data. Their axial relationships are summarized in Figure 6. Each theme and its subthemes will be described in turn.

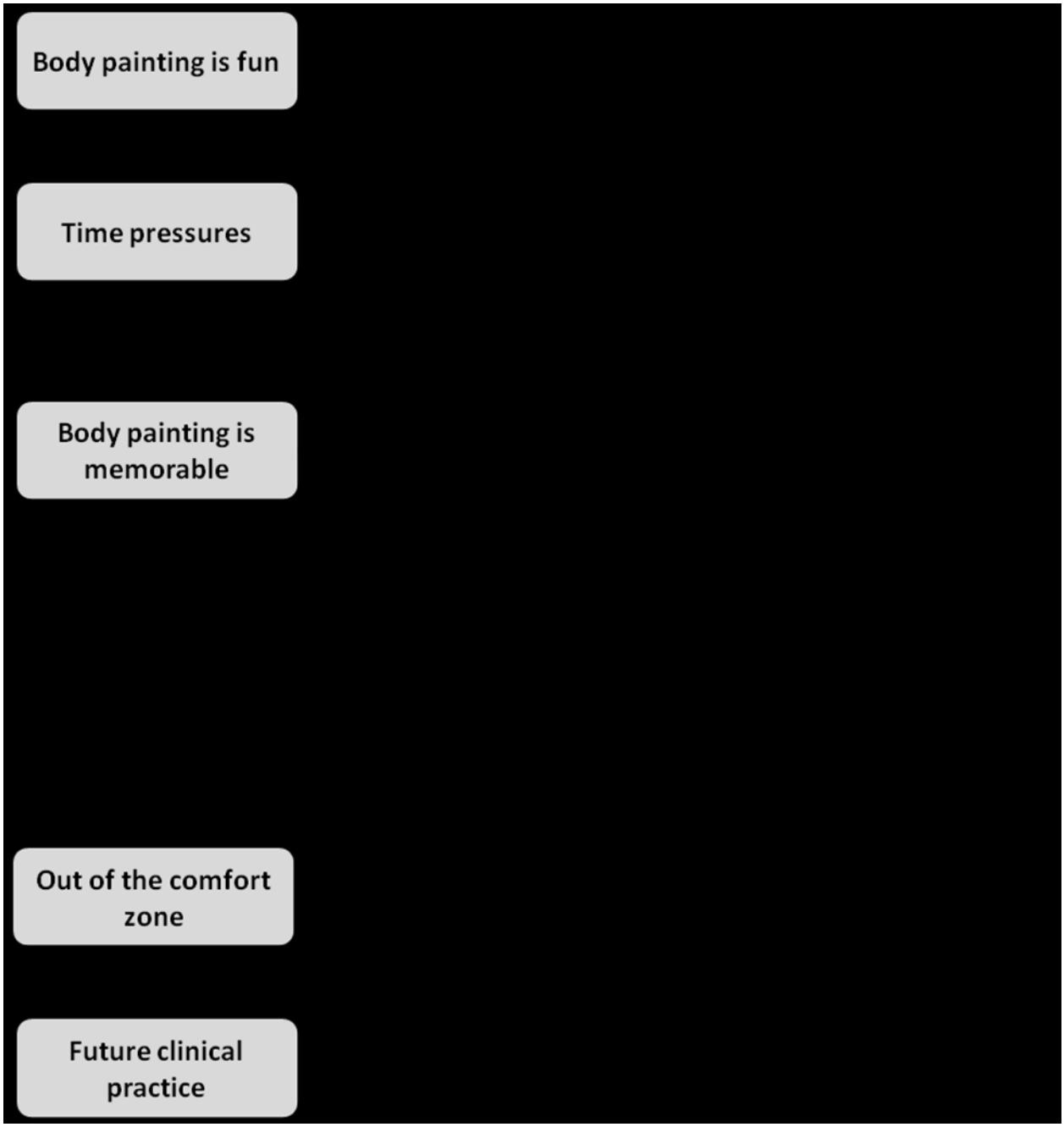

Figure 6: Relationship between themes. Themes are shown in grey. Sub-themes are shown in white.

The fun factor

The most prominent theme to emerge from the focus group data were that of body painting as a fun activity. Students used multiple adjectives to describe body painting as a positive experience; the most frequent descriptions were ‘fun’ and ‘enjoyable’. Students proceeded to describe how the enjoyment defused the academic setting. Statements were indicative that this fun activity and relaxed atmosphere promoted learning.

“...It’s fun as well, the group gets together and has a good time with body painting and that’s really, that accelerates learning a lot as well because you don’t tend to forget

that.”

“Body painting is fun, very enjoyable and often a challenge. You do not need to be a good artist, just to be able to follow instructions. Also body painting allows you to realise the situation of organs in the body, often for myself I find it surprising their exact locations and actual size.”

“...I think it is fun and you don’t really realise that you’re learning something but then it kind of sticks to your mind.”

Time pressures

A number of comments related to the time available, indicating that not enough time had been allocated to some sessions.

“I think it is really good, the only thing I’ve found with the body painting ...is I always feel like I’m under a lot of time pressure...we literally in the time that we’ve got we manage to paint stuff on, not even particularly well but I haven’t had any time to look at what I’ve actually painted on and ... stand there and be like ‘oh ok, this is what this is and this is what we’re like trying to show on this’. I don’t know, I never feel like we have time for that I kind of feel we’re so rushed,

painting it on, that I really think in the body painting sessions I could do with like double the time that we actually have to get more out of it."

Why is it Memorable?

Factors impacting upon the memorability of the anatomy learnt emerged as a theme. Students offered a variety of offered explanations as to why body painting was a memorable exercise, attributing it to four key factors: body painting as a sensory stimulus, colour as a memory aid, painting as an interactive process and body paintings as visual reminders. A further sub-theme related to the long-term retention of knowledge.

Body painting as a sensory stimulus

Body painting is a tactile process. Students find that the sensation of being painted aids their memory.

"...I think it's the tactile association so I always volunteer myself to be painted on because being totally selfish I find that I learn more from that, because then I can see the areas on myself and I've also got the sensation of it being actually painted on so I can make the association there ..."

"I remember where I felt the brush."

Colour aids recall

Colour appears to have the most significant impact upon students. This is especially true when students choose to paint using a bold palette.

> *"...Being able to see in colour and envisioning where it is, I think I could still probably envision it on myself again..."*

"...I think the major impact is the colour because I'm extremely sensitive to colour and I use that a lot in learning...If you remember a bright green stomach and a bright green bit between the shoulder blades, that's the connection, the green, you know it's not the region, it's the colour and that takes over to quite a large degree..."

Visual reminders

Students often chose to leave the body paintings on their skin for the remainder of the day. They report the visual impact as a boost to their memory.

"...Because it stays on me for the rest of the day so I'm constantly reminded of it..."

The painting process

Painting appeals to students, especially kinaesthetic learners, as it is an interactive process.

"It is useful because it is a form of interactive learning and I find that I retain information much better. The process of painting is useful."

Students were divided as to whether their preference was to paint or be painted.

"I would say painting because the actual physical act of painting helps you to remember whatever you're painting. However, being painted can be good because you have to read the partner the instructions and discuss with them where they're making the marks."

Some students expressing a preference for being painted made remarks which linked to earlier comments that retention of knowledge was aided through tactile sensation, colour and visual reminders. The visual learners appear to be exhibiting preference for being the subject.

"I like being painted and then looking. I got home after referred pain and I looked in the mirror again just to remember it and look at, and I love colours, I learn in colour."

"...It's just the fun you have basically and you can see what's been painted on yourself and you know when I'm at home and I'm in the shower and I'm looking at this, I think 'oh yeah I remember that, that was fun'..."

Some students expressed a preference for painting a peer. This preference was described by students who employed a kinaesthetic approach to learning.

"Doing it [painting], because you're thinking about where you're going...that process of actually sticking paint on."

"*Personally I find painting on other students more useful than being painted on myself, especially the practical when we painted regions of referred pain. The actual act and thinking about what you're painting onto your fellow student reinforces the learning process as it allows me a more vivid recall method."*

Retention of knowledge

Another sub-theme described how body painting promoted long-term retention of knowledge.

"...We were painting on each other last year, and I can still remember all the sites that I painted on my body and like we're talking about a year ago..."

"You remember it so well. I could remember it weeks later, especially the bright spots."

Out of the comfort zone

A fourth theme is that a number of students reported that body painting had taken them out of their comfort zone, in terms of body image and vulnerability. However, this frequently led directly to comments relating to a fifth theme, that of future clinical practice.

Body image

Body image was described from two perspectives, that of the students and that of their prospective patient.

"*I assume that body-painting makes everyone wish they visited the gym more, it certainly does that to me! But I think this is probably quite a key concept as I'm sure a lot of our patients will be wishing something similar. It is also re-assuring to see that very much the majority of people aren't chiseled like Brad Pitt."*

"I'm quite comfortable, so maybe before it was a bit hard for me to appreciate other people feeling uncomfortable, but my partner was very unconfident so I think that helped me understand more about body issues."

Vulnerability

The students spontaneously related their vulnerability to the vulnerability of patients when being asked to undress, and had reflected upon how this will shape their behaviour and attitudes in the future.

"Body painting is great fun but it does remind you of how uncomfortable it can feel to have to get undressed in front of others. It helps us as doctors because it really does make you feel what a patient would feel in a consultation situation and how certain aspects of body painting make you feel. It also makes you realise how important it is to explain to patients what you plan to do and that they can tell you if they feel uncomfortable, etc. I think it also helps with communication as you feel often quite vulnerable when being painted so you can remember how you felt when you are in a similar clinical situation in the future to change your behaviour for the benefit of the patient."

Future clinical practice

As the comments in the 4th theme clearly foreshadow, using body painting as a learning tool encouraged students to address issues with body image and confidence, and to consider aspects of professionalism and clinical practice, for the most part from the perspective of future patients. However, the painting process can be coupled with concurrent peer-physical

examination to allow students to develop their clinical skills and gain confidence in approaching patients.

"It has definitely made me more comfortable with touching patients for clinical examinations."

"It has helped me to interact with individuals in a manner that is outside the normal socially acceptable level of conduct. This has helped me develop a more professional attitude towards patients and I see this as very useful.

4.5 Discussion

A decline in the number of cadavers available for medical education, combined with the increased expense incurred running a dissection facility, has forced many institutions to seek other means for teaching anatomy. The obvious solution to such a predicament is to increase the use of medical imaging, potted specimens and plastic models. While each of these media have educational benefits, none show living anatomy. However, body painting may be another option open to institutions.

This data has shown that body painting is an enjoyable activity for students, perhaps offering a respite from intense didactic teaching. Painting defuses the formal academic context, enabling students to relax and thus promoting a positive learning environment. It may well create 'learning landmarks' in the form of memorable experiences, which allow readier access to recall of anatomical information. It allows anatomical information to be taught using a living subject and often, its functionality to be demonstrated concurrently. Body painting is a relatively low cost medium by which living anatomy learning outcomes can be achieved.

Marking anatomy on the body, for example the boundaries of viscera on the skin, is not a new concept. McMenamin documented the use of marker pens in anatomy teaching[89]. Body painting, in our experience, appears to have a more significant impact upon the student than line drawing, due to the use of colour. Outlining, for example, viscera with marker pens will undoubtedly have educational benefits, but it is conceivable that students will soon find this monotonous and the visual impact will very quickly be lost. Students using body painting have reported colour, and the visual reminder it leaves, to be one of the major benefits of body painting, particularly for those students who are visual learners. Students can utilise an anatomically correct colour palette when painting, but our student testimonies illustrate that the bold use of colour aids students' retention of the anatomical content painted during the session. This is beneficial as knowledge of basic science has been shown to diminish as students progress through their medical training[188] .

Medical education is a process of social and personal transformation resulting in students, and clinicians, learning to care for and about patients by actively engaging with people[35] . This transformative process can be 'conducted in clinical skills centers where opportunities for integrated learning between basic and clinical science can be achieved' using the media mentioned previously[35]. Durham students perceive body painting to be an activity by which they can develop such professional behaviours and attitudes. With respect to this, a particularly striking theme to emerge from the data were the vulnerability felt by students when participating in body painting sessions. This vulnerability could be perceived to be a negative experience by participating students and faculty considering implementing painting. However, the students' comments indicate gratitude that they had been given an

opportunity to experience examination and nudity from the perspective of a patient, by being taken ‘out of their comfort zone’. By engaging students in activities which stretch their boundaries, the emotional engagement with learning maybe heightened. Students’ comments also resonate with Vygotsky’s ‘Zone of Proximal Development’ which highlights the difference between what a learner can do without help and what they can do with help. Vygotsky suggested that the role of education is to provide the learner with experiences which are in their zone of proximal development, thereby encouraging and advancing their individual learning. With regard to body painting, the students’ removal from their comfort zone encouraged them to learn about the vulnerability felt by patients. The zone of proximal development is the gap between what a learner has already mastered (the actual level of development) and what they can achieve when provided with educational support (potential development). Without faculty support and the body painting activity, students may not have placed themselves in a position whereby they felt uncomfortable, and such learning may not have occurred.

Experiencing vulnerability and helplessness, albeit at the hands of a peer, has provided students with the opportunity to mould their future clinical practice and attitudes. This finding supports McMenamin [89] who stated that body painting had “the added bonus of helping break down apprehension regarding peer–peer examination.” Stillman et al. (1978) suggest that repeated contact with a partially robed or disrobed model may then student’s comfort in actual clinical situations with patients [189]. This is also likely to be true of partially dressed peers when engaging in body painting. Body painting has also encouraged students to confront issues surrounding their own body image. This finding may provide evidence that

body painting helps reduce Objectified Body Consciousness (chapter 1), as described by Rees[77]. From the aesthetic view point, this study has found that students often spontaneously use body painting for their own study and recreational purposes. This indicates greater comfort with the body and wider appreciation of the various meanings which can be ascribed to it.

Any new teaching intervention must have an educational benefit. Students have described body painting as an activity which promotes learning in many ways. Its appeal is widespread, due to its practical nature. Body painting supports many learning styles: tactile, visual, kinaesthetic and auditory. The most prominent themes to emerge from analysis of our data were the sensory impact both painting and being painted had upon students. This was both tactile and visual in nature. Thus, living anatomy, and in particular body painting, adds a new dimension to anatomical study, often not as achievable by traditional cadaveric methods. Body painting is not suitable for all anatomy teaching, and should not be overused. Sensitive areas of the body are best avoided. However, other options including hiring life models are open to institutions if such areas were to be painted. With appropriate and selective use it can have a significant educational impact on students in terms of anatomical knowledge and professional development.

Friere (1973) suggested that the development of professional knowledge is dependent upon a connection being made between the classroom and practice knowledge in context, crucial for reducing the gap between theory and practice[190]. Body painting allows educators to link anatomical and clinical teaching, in a semi-clinical context, allowing the functional and clinical relevance of anatomy to be sign-posted to students. Educators must help students to see the

link between teaching and their future practice, since this will both capture and sustain the students' interest[179]. Students are provided with the opportunity to improve their communication skills and develop a professional approach to patients in a low stakes, relaxed environment. Following body paint application students can participate in clinical skills examinations, visualising the viscera that they are palpating simultaneously due to the bold markings on the skin. Simultaneous rehearsal of auscultation and palpation has been described previously [36, 88, 89]. Clinical skills teaching sessions are often fewer in frequency during the early years of medical training, and body painting provides an additional opportunity to rehearse clinical examinations during an early stage of undergraduate medical training. Students' abilities to relate previous knowledge with new knowledge, and to relate knowledge from different courses, are characteristic of deep learners [172]. Thus, using body painting as a tool to link clinical and anatomical aspects of the curricula ought to encourage students to employ deeper learning strategies. This adheres to the guidance from Drake that[5], regardless of the curricula structure, "it should emphasise active rather than passive learning and the development of lifelong learning and problem-solving skills while simultaneously decreasing lecture time and the need for memorisation" [5]. Furthermore, body painting promotes active learning, whereby the learners are responsible for their own learning [191]. Bruner (1961) suggested that students are more likely to recall information if they actively engage with the material [192]. Active learning promotes students to adopt deep learning approaches and to "search for an understanding of the subject rather than just reproducing what they have learned" [191].

Students alluded to the fact that their personal preference and indeed learning style reflected heavily on their preference for the role they took in the teaching sessions. The cases presented for both being painted and the painter provide educators with a strong evidence

base for the educational benefits of body painting as perceived by the students. It would therefore be our recommendation that students are encouraged to select their favoured roles, in order to gain as much as possible from this intervention.

Anatomy is considered to be a subject of high factual content, often leading to learners employing surface learning strategies. By participating in this active approach to learning, students will utilise deep learning methods, synthesising knowledge of the position of the anatomical structures in question and in addition gaining an understanding of the associated functionality and clinical skills. Successful learning is achieved when new knowledge is incorporated into current understanding, supplementing or even superseding previous knowledge[171]. It is for this reason that body painting lends itself as an adjunct to traditional teaching; knowledge acquired during didactic teaching and other learning activities can be consolidated, built upon and applied during body painting sessions. This study also indicates that body painting may contribute to what Jason (2007) calls 'emotional literacy', described as an essential for successful teaching[179].

“Emotional literacy is our ability to read and understand emotional states, and most importantly, to act on them”[193]. Emotional literacy involves sensing, understanding and using emotional information from oneself and other people in order to make informed and effective decisions; this then drives actions. Emotional literacy involves skills training as well as developing personal awareness. During activities such as body painting, students are dealing with issues including nudity, body consciousness and communication. The link between emotional literacy and clinical examination become evident when considering such issues. Engaging students in body painting and peer-physical examination could therefore

provide a safe environment in which to master emotional literacy for the clinical or practice setting.

Anatomy education is a discipline shaped by its history. As educators we have been socialised into methods of teaching which are historical and therefore often an expected tradition. With technological advances, imaging and the virtual context of teaching and learning have, quite rightly, an increasingly important role in anatomy education. It is hoped that the testimonies of our students may encourage institutions to engage their own students in body painting, and view it as a beneficial teaching tool which allows students to learn anatomy whilst simultaneously developing their professional attitudes. It is not suggested that body painting ought to be a substitute for traditional teaching, whether it be using cadavers or by other means, but the role of body painting as a motivating, fun and educationally useful adjunct to traditional anatomy and clinical skills teaching is advocated.

Part 2. Body painting and line drawing: a cross-over study.

4.6 Contextual framework

The second part of this study was informed by the qualitative study (part one[90]) exploring students' perceptions of body painting as a tool for learning anatomy. Our students strongly advocated the use of body painting in undergraduate anatomy and clinical skills teaching. Presentation of these data at conferences[85, 181, 182] raised questions from anatomists in the audience regarding the practicalities of implementing body painting, and why was it any better than line drawing. The author (GF) then began to question its measured educational impact, as opposed to its perceived benefits.

Anatomy teaching has a strong tradition of line drawing, dating back over a century[184]. Line drawing requires only the marking of borders of anatomical structures on the skin, it does not allow for the use of colour, neither can texture nor dimensionality be easily added. Students are, however, following the same process in constructing line drawings since they inspect the same bony landmarks as they would when body painting, and produce a similar, but more simple, map on the skin. It was in response to this that part 2 of this study was designed. The research questions to be answered were:

- Is body painting any better than line drawing as a tool for learning anatomy?
- Does one medium lend itself better to long term retention of knowledge?

4.7 Methodology

4.7.1 Ethics

The cross-over study was approved by the ethics committee for the School of Medicine and Health at Durham University. Evaluation of teaching did not require ethical permission.

4.7.2 Participants

Ninety-seven 1st year medical students were invited to participate in a study to evaluate the effectiveness of body painting and line drawing as tools for learning surface anatomy.

Students were recruited via a presentation at the beginning of a compulsory anatomy teaching session. Information sheets outlining sponsorship, ethical approval, data collection, data protection and dissemination were circulated. Consent forms were not completed, instead students consented during the pre-test by using their Keepad™ (explained later).

4.7.3 Study design

This study followed a cross-over design. Pre-intervention, mid-intervention (immediate post) and post-intervention testing was utilised to assess whether learning occurred and the students' level of recall. The author (GF) conducted a participant observation and colleagues completed field notes during the process.

The study design is mapped in figure 7.

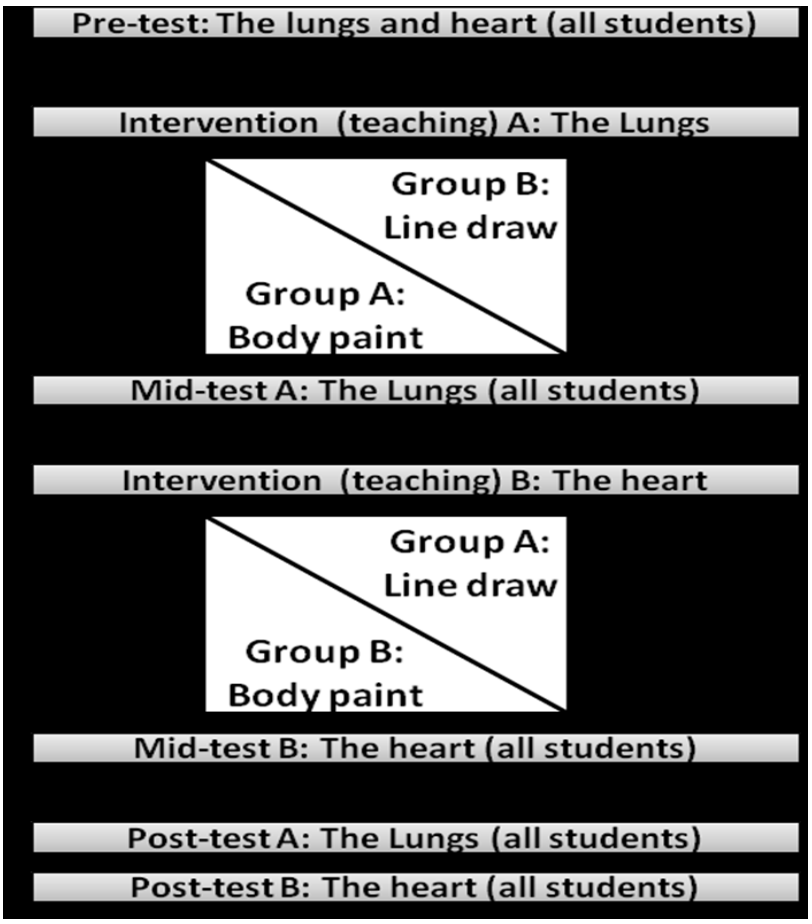

Figure 7: The cross-over design for this study.

4.7.3.1 Rationale

This study was designed as a cross-over to ensure that all participants were given the opportunity to experience both body painting (BP) and line drawing (LD).

4.7.3.2 Pre-intervention

Students sat a test (pre-test) to establish a baseline for content knowledge prior to the teaching intervention. At the beginning of each academic year students are randomly assigned anatomy groups (A1-4 & B1-4). This pre-test served to establish whether the randomly assigned group allocations were even with respect to level of knowledge prior to the teaching intervention. Any bias would need to be compensated for during data analysis. All students were issued with a Keepad™ audience response device. Keepad™ enables the user to answer questions by selecting from a series of options. For the purposes of this study, all responses were anonymous.

The test format was multiple choice (MCQ). Each test used in this study consisted of several MCQ questions, each based upon an image, the images related to surface anatomy of the body region being assessed. Students were given 5 answers from which to choose one correct answer. Tests were constructed and delivered using Microsoft Powerpoint 2007 (for Windows). The Keepad™ system uses Turning Point to track audience responses.

All tests used in this study were written by the author. Each question was then checked for face validity and each test for content validity by Dr Debra Patten, Lecturer in Anatomy at Durham University. Dr Patten was responsible for the curriculum design and therefore was able to ensure that the questions matched the desired outcomes of the teaching sessions.

The pre-test contained questions relating to surface anatomy landmarks for the heart and lungs, as these were the two regions of the thorax the teaching intervention covered. This pre-test was delivered 14 days before the first teaching intervention to ensure that prior

knowledge was being assessed. The pre-test was also used to collect consent for participation in this study. Students selected either ‘yes’ or ‘no’ to specify whether the data from their Keepad™ could be utilised by researchers.

4.7.3.3 Intervention A: The lungs and pleural fields

The teaching intervention for this study was delivered as a cross-over design, consisting of two facilitated practical sessions (intervention A & B) and two mid-tests (immediate post-tests). The first of the teaching sessions covered the anatomy of the lungs and pleural fields. Within this teaching session students rotated around 3 anatomy teaching stations in their pre-allocated anatomy groups. Stations 1 & 2, each 30 minutes in duration, were based within the dissecting room (DR), where they were taught in small groups using cadaveric prosections. Station 3 was a one hour long body painting and surface anatomy session. The intervention in this study related only to station 3. During this teaching session students were divided into their anatomy groups. As station 3 was a double session, there were two anatomy groups present at any one time. Formal teaching in station 3 was delivered to both groups simultaneously to ensure consistency. Formal teaching was provided by a clinician, who described to students the process of counting ribs and locating bony landmarks of the thorax. Students were then divided into their anatomy groups. One group was allocated to the body painting condition (BP), whilst the other was allocated to the line drawing condition (LD). The two groups, based within the clinical skills lab, were separated by a partition. Each group had its own facilitator, plus additional support staff who floated between the two groups.

Students were issued with guidelines which gave step by step instructions outlining how to draw or paint the borders of the lungs onto the thorax of a volunteer. Students self-selected a partner for this exercise. Within their pairs students either painted or drew the points listed in the guidelines onto their partner, or acted as the canvas. The decision as to who painted or drew and who was the canvas was left to students.

Durham University's stance on participation in this type of teaching activity is that it is compulsory when combined with concurrent clinical skills teaching. Students may choose which role they take, but painting a classmate is compulsory, because it requires students to rehearse locating bony landmarks, vital for future clinical practice. Being the canvas requires a level of undress and is therefore optional. Students not willing to be the canvas may also opt to assist the painter. To avoid the canvas taking a passive role they are encouraged to read the guidelines to the painter and observe the paint application as best they can.

During the painting process students are required to locate bony landmarks in order to mark the position of the underlying viscera. The points marked in each test condition (BP & LD) were identical, the only difference was the medium used.

Body painters used commercially available, water based, hypoallergenic paints, applied using a brush. Line drawers used cosmetic pencils which could be applied directly onto the skin.

After completing the teaching intervention, students were then tested. This mid-test (mid-test A) consisted of 11 questions relating only to the structures just painted, i.e. the lungs and pleural fields. Students were clothed during the testing process and were therefore unable to see their paintings; instruction sheets were also removed from view.

At the beginning of the mid-test, prior to answering the 11 anatomy specific questions, students indicated their anatomy group, the activity they had just completed (BP or LD) and the role they had undertaken: canvas, painter or helper. These answers were used during data analysis. Mid-tests were utilised to ascertain whether learning had occurred.

4.7.3.4 Intervention B: The heart and its valves

This teaching session was delivered 2 weeks after intervention A. The procedure for intervention B mirrored that of intervention A. This time students painted or drew the borders of the heart and the positions of valves. Test conditions (BP and LD) were cross-over, therefore students who painted the lungs previously now drew the heart and vice versa. This is demonstrated in figure 7.

Immediately after this teaching session, students completed a heart specific mid-test (mid-test B) which tested the content of the teaching session. The format was again eleven MCQ questions, as well as students indicating their test condition, anatomy group and role using Keepad™.

4.7.3.5 Post-test and evaluation

Twelve weeks after intervention A, and ten weeks after intervention B students sat two post-tests. Post-test A was a repeat of mid-test A, and post-test B was a repeat of mid-test B. The post-tests were given to students in order to assess whether or not there was any difference in the retention of knowledge for students taught in each condition (BP or LD). For both of the post-tests students were asked to identify their anatomy group, their role and their intervention condition for each of the relevant teaching sessions (intervention A & B). Having both anatomy group and intervention conditions allowed the author to ensure that test scores were analysed correctly with respect to teaching method (BP or LD).

4.7.3.6 Evaluation questionnaire

Students were also issued with a questionnaire at the end of the final teaching session. Questionnaires were distributed to all students, who for the most part filled them out immediately. This sought to evaluate the teaching session. Students were asked a series of questions and were given the opportunity to write free-text comments about their experiences of both line drawing and body painting. This questionnaire evaluated the teaching sessions and was therefore routine and thus was not part of the ethics application for this study. The decision to issue a questionnaire was informed by the author's participant observation. Students' comments regarding their preference for body painting and line drawing could not be quantified by the results of the cross-over study, nor could their comments to the researcher be accurately recorded, hence the inclusion of the opportunity for free-text comments.

Content analysis was performed on free-text comments, using a grounded theory approach, the methodology for which is described in Chapter 2. These findings will be reported later in this Chapter.

Quantitative questionnaire data were tabulated using Excel (2003 for Windows). Questionnaire data were analysed using simple descriptive statistics (percentage scores,%) in order to ascertain students' preferences (Microsoft Excel for Windows). Results were then cross-tabulated (SPSS version 16.0 and Microsoft Excel for Windows). Cross-tabulation was based upon students' sex (the influencing variable) and the questionnaire item (the dependent variable); for example, were males (the influencing variable) more likely to prefer body painting (the dependent variable) compared to females (the influencing variable)?

Following cross-tabulation, frequencies were produced (SPSS version 16.0 and Microsoft Excel for Windows); analysis of frequencies, independent and dependent variables was conducted using Chi-squared (SPSS version 16.0 for Windows). Chi-squared was chosen as the data were frequencies, and nominal, and Chi-squared does not impose a limit on the number of independent categories allowed.

4.7.3.7 Participant observation

The author acted as a participant observer[153], instead of a facilitator for the practical aspects of this study (interventions A & B). During this observation the author immersed

herself in the activity, by observing both activities and conversing with students. The purpose of this observation was to evaluate the teaching session, in particular the practicalities of using the two media (BP & LD). As the observation was part of a teaching evaluation, it was not part of the ethics application for this study. It was emphasised to students that the participant observer was interested only in how the sessions were running. Field notes were not made during the teaching session as it was felt that this may distract students from the activity, instead the author made field notes immediately after each teaching session. These notes were written in continuous prose using Microsoft Word 2007 for Windows.

The author's field notes focused upon student engagement and process practicalities. During conversations with the students during the session the author was particularly attentive towards the issues students were encountering whilst participating in the learning activities (BP & LD). These conversations allowed confirmation of the author's perceptions of the issues faced. This was a constructivist approach, as the new information gained shaped the remainder of the data collection. A consequence of this was the introduction of the evaluation questionnaire already described.

Staff who had a facilitatory role in the intervention sessions (both A& B) also completed field notes outside of the teaching session and emailed them to the author. Field notes were made by staff in any way that they saw fit. No instructions were issued to staff, they received only an email invitation to report back on their observations. Staff were made aware that their notes would be triangulated with the focus group data from part one of this study. Staff were notified that emailing notes to the author was indicative of their informed consent for the use of this data. Staff routinely make field notes after a teaching session,

this is in order to inform future practice. For this reason, these field notes were not part of the ethics application for this study. After the questionnaire was issued, staff permissions were requested to triangulate their comments with the student views.

The purpose of including staff field notes was to answer the following research question:

- What do staff perceive to be important in ensuring a successful body painting teaching session?

The rationale for further comparison of the staff notes with the subsequent questionnaire comments was to answer the following questions:

- Are staff and student perceptions of body painting and line drawing the same?
- With regard to the success of a teaching session, are staff and students influenced by the same factors?
    - If so, what are these factors?

Field notes were compiled (Microsoft Word 2007 for Windows). NVivo was not utilised as field notes were not extensive. Field notes were coded using the same method of content analysis, grounded theory, as described in Chapter 2. Student comments (themes) were then triangulated with staff feedback on the teaching session, as recorded in their field notes. These themes, and their exemplary quotes, were then inspected to identify any similarities with or differences between staff and student views. The student views were those expressed in part one of this study[90] and the free-text questionnaire data.

The qualitative data were further substantiated by triangulation with the quantitative questionnaire data, and the quantitative data from the cross-over study.

4.7.3.8 Triangulation

Barbour [194] suggests that, “triangulation addresses the issue of internal validity by using more than one method of data collection to answer a research question”[194]. Arguments against triangulation in qualitative research are focused upon the differences in approaches. Different data collection methods, for example interviews, field notes and focus groups, are not readily comparable. Each method provides only one viewpoint on the research question. Obtaining similar findings from different methods may provide corroboration or reassurance; caution should be aired when like findings are absent. Barbour advocates that such absence does not provide grounds for refutation.

Exploration of a multitude of viewpoints has been, within the qualitative paradigm, referred to as crystallisation [194-196]. Crystallisation “allows for infinite variety of angles of approach[196].” Much debate persists concerning triangulation and crystallisation. Richardson (2000) argued that the world has “far more than three sides”[195]. It was this argument that led to the concept of crystallisation, whereby the geometry of a crystal, with its multiple sides, can be compared to the multiple viewpoints on any one topic. However, Tobin and Begley (2004) warn researchers that crystallisation is not well documented in the literature, and runs the risk of reinventing the triangulation wheel[196].

In light of the literature, the qualitative results from the evaluation questionnaire were, with caution, briefly compared to the results of the focus group study [90]. “Triangulation and crystallisation are means of establishing completeness in naturalistic inquiry, rather than confirmation[196].” The purpose of this comparison was not to internally validate findings but, to look for relationships and consistencies in results of these two studies. The two studies consider the viewpoints of different cohorts. By comparing the data sets, it was hoped that key themes would emerge, thus, by considering multiple viewpoints, identifying the factors which students deemed to be fundamental to the success of this teaching intervention.

Triangulation is seen by some as a contentious issue, especially from the perspective of a purist constructivist. I would argue its worth in enabling the researcher to confirm their findings. For example in this study, quantitative methods allow me to identify if a difference exists between the conditions, whereas qualitative highlight the factors which impact on the conditions. Furthermore, triangulating data from different sources strengthens this relationship. As long as the researcher is aware of the limitations of each methodology and that the data has arisen from different sources, I believe that triangulation can only serve to strengthen the argument a researcher presents.

4.8 Results

4.8.1 Quantitative results

Of the 97 invited participants, 93 consented for their data to be included in this study (96% participation). From the data collected, 82% were included in the analysis. Data were

excluded if participants had not registered which activity (BP or LD) they had participated in for each part of the study (5%),if they were absent from one of the intervention teaching sessions or tests (6%), or if their Keepad™ responses failed to register (7%).

For each set of test data (pre-, mid- and post-) a Kolmogorov-Smirnov Test of Normality was performed to ascertain whether data followed a normal distribution (SPSS version 15.0 for Windows). All data were normally distributed and could therefore be analysed using an independent t-test.

Pre-test data for the groups assigned to each condition were analysed using a t-test. The independent t-test was utilised because the variances for the groups were found to be equal, ascertained using an F-test (F=0.431, p= 0.536). The mean scores for the groups were not significantly different (t= -0.003, p=0.998). It could therefore be assumed that the distribution of students between groups was equal, with respect to prior content knowledge.

Mid-test data for both intervention A & B were analysed, collectively, with respect to the activity undertaken (BP or LD) and the role undertaken (painter or canvas). Considering the activity undertaken, the F-test ascertained equal variances (F= 1.614, p=0.206). The mean scores for body painters and line drawers were 53.05% and 56.99% respectively. This result was not significant (t=-1.449, p=0.149).

After determining equal variances (F= 3.456, p=0.065) an independent t-test found the mid-test scores were significantly different when considering the roles students undertook (t=2.069, p=0.040). Those acting as the painter/ line drawer had a mean score of 57.65%, whereas the students whom were the canvas scored 52.05%.

Post-test data, again with equal variances for activity (F=0.792, p=0.375) and role (F= 2.167, p= 0.143) found there to be no significant differences between painters and line

drawers, mean scores of 61.38% and 62.57% respectively (t=-0.564, p=0.574), and between the painter/ line drawers and the canvases, means scores of 62.25% and 61.60% respectively (t=0.289, p= 0.773). All mid and post test scores are summarised in tables 6 and 7 respectively.

	Mid-test	
	Percentage score (%)	p-value
Body painters	53.05	0.206
Line drawers	56.99	
Painters/ drawers	57.65	0.040
Line drawers	62.05	

Table 6: Mid test scores for each experimental condition

With respect to retention of knowledge, students' mean mid-test and post-test scores, for all students in both conditions (BP & LD) combined, were 55.02% (mid-test) and 61.98% (post-test) respectively. These results appear to show and increase in knowledge recalled between the post and mid tests. A one-sample t-test (SPSS 15.0 for Windows) was performed to ascertain whether or not these mean scores were significantly different from each other. The increase in knowledge from mid-test to post-test was statistically significant ($t=16.81$, $p=0.038$). The effect size was calculated and deemed to be 'large'[197] (cohen's $d = 33.62$, $r= 0.998$) .

	Post-test	
	Percentage score (%)	p-value
Body painters	61.38	0.574
Line drawers	62.57	
Painters/ drawers	62.25	0.773
Line drawers	61.60	

Table 7: Summary of mean post- test scores (%) and p-values. * denotes a significant result.

Of the 93 participating students, 84 returned evaluation questionnaires (90.3%).

Question / Percentage (%) score	Body painting	Line drawing	Neither
Ease of application	39.3	54.8	4.8
Educationally beneficial	56.0	33.3	9.5
Enjoyable	76.2	14.3	8.3
Surface anatomy	46.4	40.5	11.9

	Yes	No	Don't know
PPE barriers	77.4	8.3	13.1
Recommend	86.9	2.4	9.5

Table 8: Scores for BP and LD from evaluation questionnaire

4.8.2 Qualitative results

4.8.2.1 Qualitative results from the evaluation questionnaire

Content analysis of the qualitative free-text data from the evaluation questionnaire issued gave rise to 5 main themes: the practicalities of a session, the education impact of a session, enjoyment, roles, and staff involvement in the session.

Each theme, and its associated sub-themes, will be reported in turn; exemplary quotes have been provided for each theme.

The practicalities of a session

Students' comments within this theme were associated with the environment and the procedure. The sub-themes were: ease of application, instructions issued, mirrors, temperature and timings.

Ease of application

Line drawing was described as an easier process than body painting, mostly due to the time it took to apply.

"...line drawing....was quicker so was preferred and allowed tasks to be completed..."

"Line drawing is easier to apply and accurate, but harder to wash off."

Instructions issued

Some requests were made for instruction sheets to be more detailed.

“More concise instructions with maybe step by step diagrams rather than lengthy instructions with a random picture.”

Mirrors

Students requested that mirrors be provided.

“Maybe mirrors would be helpful?”

Temperature

Student comfort, in particular room temperature, influences students’ willingness to engage.

“It is cold. I would like better heating in the rooms.”

Timings

Timing was a major issue for students; numerous comments focused upon students feeling rushed.

“There was not enough time to fully consolidate it.”

“I think it [line drawing] is helpful but we need more time to draw.”

“...also felt rushed at the end, so couldn’t observe my beautiful artwork!”

The education impact of a session

Many comments expressed opinions on the perceived educational value of the teaching sessions. This theme reports on sub-themes of: the impact of body painting, the impact of line drawing, aids to surface anatomy, aids to clinical skills and physical examination. Within the sub-theme of the impact of body painting, a mini-theme, the impact of colour, is reported.

The impact of body painting

Students describe body painting as being beneficial.

“The body painting was useful.”

“Body painting was better.”

“Found body painting really helpful. Really good idea.”

The impact of colour

Colour appeared to have a positive impact upon student learning.

“Paint can be colour coded. Fun.”

“Body painting was better as the colours made it easier to visualise.”

“I found both useful for learning surface anatomy especially body painting as it can be colour coded and therefore easier to remember and learn.”

“Body painting was messy, annoying and it was nearly impossible to make clear outlines so organs looked like amorphous blobs.”

The impact of line drawing

Students found line drawing to be a more precise and manageable media.

“Line drawing was easier, clearer and easier to clean up afterwards, not to mention quicker.”

“Line drawing was more accurate.”

“I enjoyed both line drawing and body painting, however with line drawing being “less messy” scored higher in my opinion.”

Aid for surface anatomy

Both modalities were perceived to be a useful method by which to learn surface anatomy.

“...Useful to remember the surface landmarks and size/ orientation of organs.”

“I found line drawing benefitted my understanding of surface anatomy, but body painting was even better.”

“Found it good to learn surface anatomy and remembered it when revising.”

Aids to clinical skills and physical examination

The process enabled students to build confidence in approaching clinical examinations.

“Both sessions improved my confidence in peer examination.”

“The process made me more comfortable examining anatomical landmarks and conducting physical examinations.”

“It helped put people at ease with physical examination and also was an entertaining way to learn body landmarks.”

“Helped to overcome stress/worry of physical examination, highly entertaining.”

Enjoyment

The next theme to emerge from the data were enjoyment. Students described enjoyment and fun with respect to both line drawing and body painting.

“Fun and easy way to learn anatomy.”

“I found it fun and enjoyable.”

“Great fun and social activity.”

“Fun and good way to learn. Also helps you get to know people on the course better.”

Roles

When describing their perceptions of the impact of the roles undertaken in a teaching session students expressed preference for being the painter/ drawer.

"I think it was more useful if you were the painter/ line drawer."

"I fell I learn much more if I am the painter rather than the canvas. It is difficult to learn if you can't see what is being drawn."

"...it's not very beneficial if you don't do the painting /line drawing, but being the canvas is better for peer-physical examination."

"Sometimes ran out of time, which meant only one person could be the model/drawer/painter. If the sessions were either longer, organised better think we would have had chance to exchange roles. I think you definitely get more out of it being either the drawer or the painter."

"Beneficial when being the painter but don't find it helps much to be painted. Need enough time to both have a go."

Staff involvement in the session

The final major theme to emerge from the data were the level of involvement students required of staff. The sub-themes reported are feedback and requests for demonstrations.

Feedback

Comments indicated the students' desire for more reassurance and feedback during the teaching session.

"More time to answer feedback questions."

“Helpful- useful- although needs great moderation to check where analysing correctly as otherwise feel blind.”

“More help sometimes.”

Demonstrations

Students requested staff to demonstrate the technique at the beginning of the session.

“Also it may be useful to have fully painted models as examples.”

“Demonstration beforehand and more help.”

“Give demonstration.”

4.8.2.2 Triangulation of focus group data and questionnaire data

As previously reported (part one), focus group data resulted in the emergence of five principal themes, each with their associated sub-themes. These results from part one were:

(1) Body painting as a fun learning activity,

(2) Body painting promoting retention of knowledge:

a. Time pressures

(3) Factors contributing to the memorability of body painting:

a. Body painting as a sensory stimulus

b. Colour aids recall

c. Visual reminders

d. The painting process (role preferences)

(4) Removal from comfort zone:

a. Body image

b. Vulnerability

(5) The impact of body painting on students' future clinical practice (clinical examination).

Superficial inspection of the themes highlighted consistent viewpoints between focus group data (part one) and students' evaluation data (part two). These were that body painting is a fun activity but its success as an educational intervention is dependent upon adequate time allocation. Students are receptive to the use of colour, and have strong but divergent opinions on which role is most beneficial, that of the painter or the canvas. The final similarity which can be noted is that of the positive impact of body painting on students' confidence in engaging in clinical examination.

4.8.2.3 Participant observations

In order to truly understand both body painting and line drawing as teaching tools a different viewpoint was needed, that of the instructor. As previously described, staff routinely involved in the facilitation of body painting sessions provided field notes taken during teaching sessions. These field notes have been analysed and are now reported.

Out of the 9 members of staff invited to provide field notes taken during the teaching sessions (intervention A & B), 6 returned notes. These notes were provided by the following:

- The author of this thesis (principal researcher)
- Lecturer in Anatomy
- Technician 1
- Technician 2
- Teaching Fellow in Anatomy
- Visiting Clinical Facilitator (Paediatric Registrar)

Field notes were complied (Microsoft Word 2007 for Windows). NVivo was not utilised as filed notes were not extensive.

Coding of field notes resulted in the following themes emerging: environment, media, time pressures and instruction.

Environment

The first theme to emerge from the data were focused upon the teaching and learning environment. Sub-themes were space and mirrors.

Space

From the staff perspective, provision of ample space was a key issue.

"Students very close together, easily side-tracked."

“Room set up - xxx had students in a very small area compare[d] to xxx, closer supervision of students, less opportunity for distractions in xxx's group which [was] widely spread out.”

“It would be better if we could find some way of setting up xxx so there is enough room for everyone.”

“It would also help if we had a designated area for body painting that was warm and had cubicles.”

Mirrors

Staff requested mirrors be provided for future sessions.

“Mirrors need to be used for students to see what’s been painted on the skin surface.”

“I think we definitely need some body length mirrors for the students, it would be helpful for them to see what has been painted on them.”

“The students could do with mirrors, they would really appreciate being able to see the results.”

Problems with media

Staff described the issues they had encountered whilst working with both body paints and pencils.

Pencils

"While the eyeliners are good to use for line drawing they don't last very long and sharpening them can make the wooden edges quite jagged and nobody wants a splinter! They are also expensive so I would suggest finding another way of doing line drawing such as damp charcoal or body paints but in a soft crayon form."

"The eyeliners were effective, but very soft for drawing, there were lots of broken points and pencil shavings around – and they're very expensive!"

"The sharpeners were not good at sharpening pencils and the pencils got blunt quickly so needed sharpening a lot and was a pain!!!."

"Line drawing is so much quicker. It doesn't involve any messing about with water, brushes or changes of colour."

"...my opinion is that line drawing is actually better, in the sense that students seem to get through more and seem to have more time to digest the information as opposed to rushing through the painting."

Body paints

"Body paints get very messy. Students smudge all the colours together. It's a lot of work trying to refresh the water and clean the brushes."

"Students use too much water, need to keep paint thick for easier application and for accuracy."

"…with the paints maybe softer, more flexible brushes are needed so the edges of what is being painted can be more defined, I think the brushes we have are to hard and can get quite tough when the body paints dry on them."

Time pressures

Staff were aware of the time pressures students faced in completing the activities.

"Smaller groups and a little more time would be ideal."

"The ideal BP session would be if there was time to teach and practice the relevant clinical skill at the end of the BP session, since this is the point of learning surface anatomy. This would need careful timetabling, involving xxx (Clinical skills lead), if she could teach the clinical skill at the end and there was time for students to practice it that would be ideal ; it's what was always planned in these sessions in the beginning but lack of time."

"Students seem to want more time so that they can both paint and be painted."

"The body painters also take a lot longer to get going in the session. The line drawers can get going a lot quicker on the task and thus finish a lot sooner. The body painters seem to take a lot of time to choose the colours they are going to use whereas the line drawers have only the pencils to choose from."

Instruction

Finally, staff commented on the level of instruction given to students and the impact this may have had on teaching.

“Teaching styles - A's and B's differed, B had a more facilitatory role than A who did more teaching.”

“Tutor differences - A is a clinician, it may be that his/her group responded differently to him (see him as a role model since he/she’s qualified) compared a member of the anatomy team.”

“Teaching style- a demo of body painting / line drawing at the beginning of the session- or in a lecture theatre before hand?- then tutors can just facilitate.”

“Students expect to be taught. These activities lend themselves to peer-led or self-directed learning. Students expect staff to be on hand to confirm the bony landmarks, otherwise they appear panicked.”

“C was in a small (more “cosy” area of the lab. It was easier for the students to ask C to help them. Because of this – the students were given almost one-to-one tuition on where to draw the lines for the outlines of the organs. In the “body painting” side –our students were more spread out and we did only facilitate their progress. This could have a bearing on student satisfaction.”

4.8.3.4 Triangulation of staff and student views

Staff views appeared to be focused on the negative, or poorly executed aspects of the teaching session. This may be due to their training as reflective practitioners, for example during study for teaching qualifications where faculty are encouraged to reflect on their

teaching and continuously self-evaluate and criticise their performance in order to improve. The consistencies in opinions from staff and students were the timing issues, both concluding that it would be ideal to allocate more time for these sessions. Furthermore, there was a consensus that mirrors would aid the learning experience.

Differences surround staff involvement in the sessions. Both staff and students commented on this topic; however, the focus of each group differed significantly. Students wanted more staff involvement; this was supported by staff, who spontaneously commented that students did not get as much instruction as perhaps was needed due to the layout of the classroom. However, staff thought students' engagement may be influenced by the background of the facilitator, focusing on clinical versus academic training. Both staff and students commented that a demonstration would be useful.

With regard to the different media used, both groups found the pencils used for line drawing an easier media to work with. The reasons for this were the ease of application, thus saving time, and there was less mess than that created using paint. Staff, understandably, were concerned with the cost implications as the cosmetic pencils were expensive. Cost was unlikely to be a concern to students, as they are the consumer within the context, whereas staff assume the role of service provider.

4.9 Discussion

The results of the evaluation questionnaire clearly show students enjoy engaging with body painting as a learning activity; 77.1% preferred body painting, compared to 14.5% for line drawing, and 8.4% having no preference. The qualitative data from focus groups and questionnaires substantiates this, suggesting that this is due to the fun and social nature of the activity. As speculated before, perhaps this may also be in part due to the practical nature of the learning activity, which offers students a break from didactic teaching.

Despite the fact that there was no significant difference between body painting and line drawing in the cross-over results, questionnaire respondents reported body painting (56.6%) as being of superior educational value when compared to line drawing (33.7%), while 9.6% of students did not perceive one media to be more beneficial than the other. From the student comments the use of colour associated with body painting appears to be the major factor attributing to this. Further qualitative research, such as focus groups, may be able to explore this further.

Body painting, as evidenced by the qualitative data, is perceived by students to be an enjoyable activity; 77.1% of respondents found body painting to be more enjoyable than line drawing (14.5%). Again, some students had no preference for activity, with 8.4% reporting neither to be more enjoyable.

Body painting and line drawing were implemented into the curriculum as methods of teaching clinical skills and surface anatomy. Body painting was deemed to be slightly more effective (47.0%) as a method for learning surface anatomy in comparison with line drawing

(41.0%) by responding students. Other respondents could not distinguish between the methods as a learning tool (12.0%).

In support of our qualitative findings, 78.3% of responding students thought that using body painting and line drawing helped break down barriers to peer-physical examination. Some students did not find the activities helped (8.4%) and the remainder were undecided (13.3%).

Finally, students were asked whether or not they would recommend these techniques to other medical schools. A resounding 88.0% of students would recommend their introduction into the curriculum; 2.4% would not and 9.6% were unable to decide.

The self-report questionnaire data is useful for gaining a quick overview of the participants' views on the educational interventions. However, this data cannot be relied upon as the questionnaire was issued immediately after the session. As mentioned previously, this means responses could be subject to knee-jerk reactions or participants may have felt pressured to respond in a particular direction due to peers or the staff presence when filling out forms. I have used this data with caution, only to add further viewpoints which may not have become evident by other methods.

With regard to the cross-over study results, mid-tests did not reveal any significant difference between scores attained by those body painting and those line drawing

(p=0.149). However, those who had undertaken the role of the painter recalled significantly more (p=0.04) information than those who acted as the canvas. This result was expected because the painter was the most engaged partner in the process. The painter's role was the most interactive. The canvas can see the instructions but is often, unless mirrors are provided, unable to visualise both the painting process and the finished product. This passive role was likely to be the reason for the lower mean score for students who acted as the canvas.

The post-tests,12 weeks after the sessions, did not reveal any differences between those acting as the painter and those acting as the canvas (p=0.773). Furthermore, there was no difference between the two teaching methods studied, body paints and line drawing (p=0.574).

Direct comparison of body painting and line drawing, from mid and post-test data, revealed that neither method was a more superior tool for learning the surface anatomy of the thorax and boundaries of thoracic viscera. Neither method facilitated a higher level of retention of knowledge than the other. This result was not surprising since the activity was almost identical, both groups marked outlines. They only difference was the media used, pencils versus paints. This semblance in activity was not planned; but in practice, body painting students opted to use black lines instead of colour when painting, making the activities difficult to tell apart. This highlights a difficulty in controlling experimental conditions in a real world setting.

In order to see the benefits of using body paints, which from student testimony appears to be the use of bold colours, large areas of painting may be needed. The author and colleagues believe that body paint, as a media, is most useful for painting block areas, for example viscera and dermatomes. Boundaries, as resulted in this study, may have been too similar for any differences between the methods to be highlighted. Students choosing not to use block colour is likely to be as a result of the time pressures they reported in the evaluation questionnaire.

The participating students in this study were first years, with little experience of clinical examination and body painting. This meant that the rib counting and location of other bony landmarks was a slow process, students appeared tentative when applying the points required on the skin. It may have been better to have run this study later in the term when students were better rehearsed at locating bony landmarks. However, timetable restrictions prevented this. Again, an issue associated with educational research in the real world setting.

Reasons for the similarity in mean test scores between the two groups (BP & LD) may have been familiarity with the questions. With a test of only 11 questions it is conceivable that students recalled the questions. Furthermore, the post-test was after the formal examination period, this was to ensure retention of knowledge was tested, however may have been counterproductive as students would have recently revised the content.

The tests used in this study were oriented towards surface anatomy, which was not examined in the January exams. Surface anatomy forms part of the May Objective Structured Clinical Examinations (OSCEs). Despite this, students will have revised this surface anatomy as part of their study of the thorax in preparation for the exams.

Another factor to impact on the test scores was discussion. When completing the post-test students were seated in a lecture theatre. Despite strict instructions not to confer, students discussed answers. As students were seated with friends, and not in anatomy groups, this meant there was cross-talk from students in different intervention conditions. The level of conferring was so severe that two senior members of staff attempted to intervene.

Additionally, students would have been able to see the response of the person next to them if they so wished. Within a lecture theatre setting it is easy to see the Keepad™ of the person either side of you, this may even happen involuntarily. The conferring and observation of other people's responses is likely to have strongly impacted upon the results of this post-test.

This may, in part, be due to the exam revision period whereby all students would have studied the content recently, thus negating the impact of the intervention. The previous discussion regarding collusion between peers during the post-test may also be a valid consideration here. Students' mean mid-test and post-test scores, for both conditions (BP & LD) combined, were 55.02% and 61.98% respectively. These results, as reported earlier, show a statistically significant increase in knowledge recalled ($p=0.038$) from mid to post-

test. In addition to the timing of the post-test, immediately after the formal exam period, and the possibility of collusion between peers, the curriculum structure may account for this increase. Students are likely to have had additional teaching and learning experiences over the 12 week post-intervention period. This is due to the integrated curriculum. Content covered in the teaching session and test could have been repeated in physiology, clinical skills teaching, or subsequent anatomy teaching.

The utility, or usefulness, of an assessment has been defined as a product of its reliability, validity, cost-effectiveness, acceptability and educational impact[150]. Acceptance is an important consideration for any type of assessment, whether it be a formal exam or a smaller scale test, as in this case. Students enjoy using Keepad™, as it is both novel and easy to use. Irrespective of this, future tests may need to be administered either on paper or using the University intranet, as these make test conditions more manageable.

Another difficulty associated with educational research is that of accounting for variable student motivation and learning. Some students will be active in their approach to learning through the year, whilst others will condense study into the days surrounding the exam period.

Both parts of this study enable conclusions to be drawn about the factors influencing the success of a teaching intervention. These factors are specific to the context of this study, but may provide lessons which are generalisable.

The student and staff comments regarding mirrors, temperature and space, demonstrate the importance of the teaching environment. This is especially important in the body painting and clinical examination contexts as the environment is hugely influential with respect to students' willingness to participate and, more specifically, undress.

Although comments on staff input and feedback were varied, this is an area which requires consideration. Students have a need for reassurance and feedback. This may not always be possible on an individualised basis due to time constraints and the student to staff ratio. Students' requests for a demonstration of the practical component in this study may be a method which institutions can utilise instead. Although not feedback, this preliminary demonstration would illustrate to students what was required of them, thus providing reassurance and showing that the outcome is attainable. A demonstration had been planned for each of the intervention teaching sessions, but due to a breakdown in communication between the internal and external facilitators this did not go ahead. The demonstration would have shown students how long to spend on each part of the task, and may have made them less tentative in their approach, as well as highlighting where to mark on the skin.

The requests for a demonstration, coupled with the requests for clear instruction sheets, highlight the need for the desired outcomes of any teaching session to be explicit. It is well known that assessment drives learning[171], however, the student comments are indicative that so too does a goal. In this context, that would have been producing a painting that

matched that shown in the demonstration. This may have acted as a focus for students, ensuring that they kept to task during the painting process, which by their own admission was a fun and sociable activity. Ensuring that the desired outcomes of any teaching session are explicit would certainly be a generalisable finding.

With regard to the use of art based approaches, such as body painting and line drawing, their success appears to be founded within their interactivity. Attention, within the lecture, wanes after approximately 17 minutes[171] . The approaches in this study encourage students to be active learners. The organisation of the session is such that students are constantly changing activity, including listening to didactic teaching, reading instructions, painting and rehearsing a clinical examination. This not only appears to keep the students' attention but is, as reported in part one, appealing to all learning styles; visual, auditory and kinaesthetic learners[90]. This change of activity is not always practical, but could be something for academics to think about further, as a means by which student attention and motivation can be maximised.

This study may provide some evidence as to the effectiveness of teaching living anatomy in an undergraduate medical curriculum. Living anatomy is the basis of clinical examination. For this reason it is the connection between gross anatomy and clinical practice[198]. Ganguly and Chan (2008) believe that living anatomy cannot replace cadaveric anatomy in medical education[199]. They defend the inclusion of cadaveric anatomy in the curriculum on the grounds that students require a 3-dimensional anatomical knowledge before they can benefit from studying living anatomy.

The qualitative findings in both parts of this study provided a plethora of evidence regarding why students believe body painting and line drawing are useful approaches for introducing clinical examination. Inclusion of clinical examination needs to be carefully planned due to cultural and social sensitivities[76, 198]; the methods in this paper may provide an approach for introducing an examination which is both fun and sociable but remains educationally beneficial. Exposure of the body carries with it specific meaning[76]. Art based approaches breakdown the apprehension students may experience when exposing their bodies. The media used, from the author's observations, act as a distracter. Anecdotally, students have reported that the paint feels like their clothes and they forget they are naked. This means that once students progress to the subsequent clinical examination, the barriers to nudity have already been partially, if not fully, broken down.

McLachlan *et al* (2010) discuss peer-physical examination (PPE), stating that students ought to be able to select their PPE environment where possible, and that this should include selecting whether to work with friends or unknown classmates[200]. This emulates the body painting set-up; students at Durham are encouraged to choose who they work with and whether or not they work in the open plan or enclosed areas of the room. This is accepted practice because, as McLachlan *et al* (2010) suggest, there is not a single successful strategy which promotes willingness to participate in PPE.

McLachlan and colleagues reported students' perceiving PPE as being a valuable experience in terms of clinical skills and future clinical practice[200]. This mirrors the

findings of both parts of this study, where students reported body painting as a useful precursor to clinical examination and practice. They particularly emphasised how they gained an understanding of the vulnerability their future patients will feel when requested to undress. If not for the content knowledge gained, this may be a primary reason for body painting to incorporated into the medical curricula.

Anatomy teaching in most undergraduate medical courses in the UK has moved from traditional instruction towards an integrated, clinically-based training [201]. Schools are replacing the requirement for students to learn vast amounts of factual knowledge with anatomy taught within the clinical context. Dangerfield and colleagues [201] report that such curriculum change has led to the exploration of new initiatives for the delivery of anatomy teaching. New methods are flexible, relevant and contextual.

In 1990, Cox[202] advocated that new concepts should not be integrated into the curriculum unless they are relevant and contribute to practice. In order to facilitate retrieval of knowledge, students ought to be provided with relevant clinical learning experiences; these experiences should be in context, as encountered in actual practice[203], the rationale for this being that the next time the problem is experienced the answer will be retrieved and used in context. This suggestion mirrors the work of Godden and Baddeley[204, 205], and the contextual learning study in Chapter 5 of this thesis. Rolfe and Sanson-Fishcer (2002)[206] also campaign for contextual learning; “it is important to ensure that students are exposed to the sorts of clinical conditions they will be likely to encounter frequently in their clinical practice”.

Cost implications are a consideration when implementing any activity into the curriculum, especially given the 2010 reductions in University funding when major budget cuts were made [207]. “Clinical teaching methods should take into account a number of issues, including the amount of teaching time and other resources required[206].” Art based approaches to both anatomy and clinical skills teaching are cost effective methods of teaching large groups of students as the products, particularly the paints, are relatively cheap, little instructor input is required and large numbers of students can be taught simultaneously. Universities may need to further utilise such approaches; however, attention must be paid to ensure that the content is relevant and, as described previously, can be taught contextually.

The evaluation data from this study and the literature [206]call for further inclusion of immediate feedback into teaching sessions. Such feedback motivates students as it enables progress to be gauged with respect to factors such as knowledge, competence and faculty expectations. Faculty expectations were evidently a consideration for the respondents in this study, as was indicated by their requests for feedback and demonstrations from faculty. Feedback encourages students to adopt a deep approach to learning [90, 206]. Incorporating such feedback must therefore be a consideration when designing teaching interventions, such as body painting sessions[206].

It is well documented that medical education, and specifically anatomy education, strongly emphasises rote learning [90, 208]. Rote learning maintains students in a position of

passive learning[90, 208]; effective learning results when pre-existing knowledge is built upon, when learning is contextual[187] and when a learner has the opportunity to elaborate on new knowledge[177].

It is the role of an educator to provide teaching and learning opportunities which facilitate coherence between studies (the education context) and professional activity (the practice context). Chastonay (1996)[208] states that, “students who are provided with the opportunity to practice and apply their knowledge to the resolution of problems, have greater confidence in their professional abilities.” Furthermore, Chastonay[208] provides evidence that teaching based upon this philosophy maintains the cognitive performance of students, facilitates the development of their aptitude to solve medical problems and is perceived by students to be both motivating and stimulating.

The above statements can be evidenced by the testimonies of Durham students. The students were highly motivated by the art based approaches described; the fun and social nature of the activities engaged students immensely. The aptitude to solve medical problems is related to the link made between teaching anatomical knowledge and clinical skills at Durham. The anatomy taught facilitates the students’ learning of such skills; the students learn how to adapt their anatomical knowledge to perform a clinical examination. Finally, cognitive performance is dependent upon many factors; use of stimulating teaching and learning activities can only serve to improve performance.

Relevance is an important feature in medical education. Relevance has been described as fundamental to teaching, along with educational efficacy [208]. I would suggest, the relevance of content knowledge changes dramatically depending on the perspective you view it from. What an educator and a student perceive to be relevant can be entirely different. Anecdotally, students often disregard aspects of anatomy teaching as they do not perceive it to be relevant. Their focus is the immediate exam diet, and little weighting is placed upon future needs for such information. This can be related to propositions, the psychological term for how concepts grow and change in meaning as they become linked to other concepts[209]. For example, students are taught the position of viscera, with respect to bony landmarks, in the dissecting room. Students hang little significance on this proposition until they need to apply such knowledge in the clinical skills lab when requested to palpate the viscera. The concept has not changed, but its meaning has. The knowledge has been linked to another concept. Within any discipline it is best to begin with general concepts. These concepts later serve to anchor new knowledge. Meaningful, or deep, learning requires new information to be subsumed into concepts already known. This is important for educators to be aware of the students' perceptions of relevance and they must therefore sign-post relevance to students in order to promote their learning and motivation.

Ausubel's[210, 211] learning theory has two key principles:

1. The most general ideas of a subject should be presented first and then progressively differentiated in terms of detail and specificity.

2. Instructional materials should attempt to integrate new material with previously presented information through comparisons and cross-referencing of new and old ideas.

Body painting lends itself to this style of learning. The basic concepts of anatomy are taught within the dissecting room first then, through the use of living anatomy, they are built upon to provide the clinical information and application. Body paint provides the medium (or instructional material) that integrates the new information with the pre-existing knowledge.

It is important to acknowledge that Ausubel's theory[210, 211] was related to learning from verbal and textual sources within a school setting. However, the premise can be applied to adult learners and medical education.

4.9.1 Limitations

The limitations of the qualitative aspects of this study are that the participants are from one organisation, which means that with respect to anatomy and clinical skills teaching the students are culturally relatively homogenous.

The focus groups are of course limited by participant and researcher bias, and may have been affected by overly opinionated participants with any given group. All qualitative data analysis is subject to interpretational bias from the researcher. The reflexivity section describes the author's background and how this may have impacted upon her understanding of the data.

Part two of this study was limited by the closeness of the body painting and line drawing activities. The sample size was likely to have been too small to note any effect. Other

factors, such as timings of the post-test and collusion between peers have been described in detail within the discussion, but were likely to have impacted on the results. As described within the discussion, these factors resulted from this study being conducted within a practical and real world setting; many are unavoidable.

4.9.2 Future work

A study is planned to ascertain whether cultural differences and ethnicity impact upon a student's willingness to engage with innovative teaching methods such as body painting. A cross-over will be conducted comparing line drawing with the use of bold, block areas for body painting. It is hoped to ascertain the impact of colour on retention of knowledge.

Chapter 5: The Impact of Wearing Scrubs on Contextual Learning

5.1 Background

This study is in two parts. Part one is a quantitative study, while part two presents qualitative data. Part one of this study has been accepted for publication in Medical Teacher[212].

5.1.1 Contributions to this Chapter

For part one of this study I was responsible for the design, data collection and analysis, as well as writing the initial manuscript. Professor John McLachlan and Dr Debra Patten assisted me with the design of this study and reviewed the manuscript published in Medical Teacher[187]. Professor McLachlan advised on statistical analysis of the data.

I was solely responsible for the design, execution and reporting of part two of this study.

5. 2 Introduction

The impact of context on learning in medical education is of particular interest, since medical students generally learn in one setting (the medical school) and apply their learning in another (the practice setting). Whilst it is acknowledged that some schools now do more teaching in the clinical setting, especially during the early years, the teaching of gross

anatomy for the most part remains dissection room based. As a result students are still learning anatomy in a different setting to that in which they are expected to practice it.

The classic study of Godden and Baddeley (1975) suggested strong contextual influence on recall[205]. This finding might suggest that learning in an educational context might not transfer well to practice environments. If this were to be the case then educators may need to move towards teaching in environments simulating clinical practice in order to promote recall of knowledge once students have left the medical school (classroom) setting. There is considerable interest in the value of simulation approaches in medical training and contextual factors are clearly relevant to this.

Godden and Baddeley[205] did not experiment in a ‘real world’ setting. Although their findings were powerful, the context (recalling words during underwater submersion versus on dry land) was not something of use to the participants. The irrelevance of both their content (word lists) and context (water versus land) is likely to have had a negative impact on the levels of recall.

This study explores the impact of an authentic aspect of context (wearing scrubs) on learning and recall for medical students, using a qualitative approach.

5.3 Method

5.3.1 Ethics

Ethical approval was given by the School of Medicine and Health Ethics Sub-Committee prior to the commencement of this study.

5.3.2 Study design

A cross-over design was used to compare the effects of context upon information recall. This is demonstrated in figure 8.

All 93 first year medical students at a UK medical school were invited to participate in the wearing of scrubs during an anatomy teaching session on the kidney. Recruitment followed the procedure outlined in Chapter 2.

Participation in the teaching session associated with this study was mandatory as the teaching component was part of the anatomy curriculum. Wearing scrubs was not compulsory. Completion of online tests was also compulsory as they formed part of the timetabled anatomy teaching.

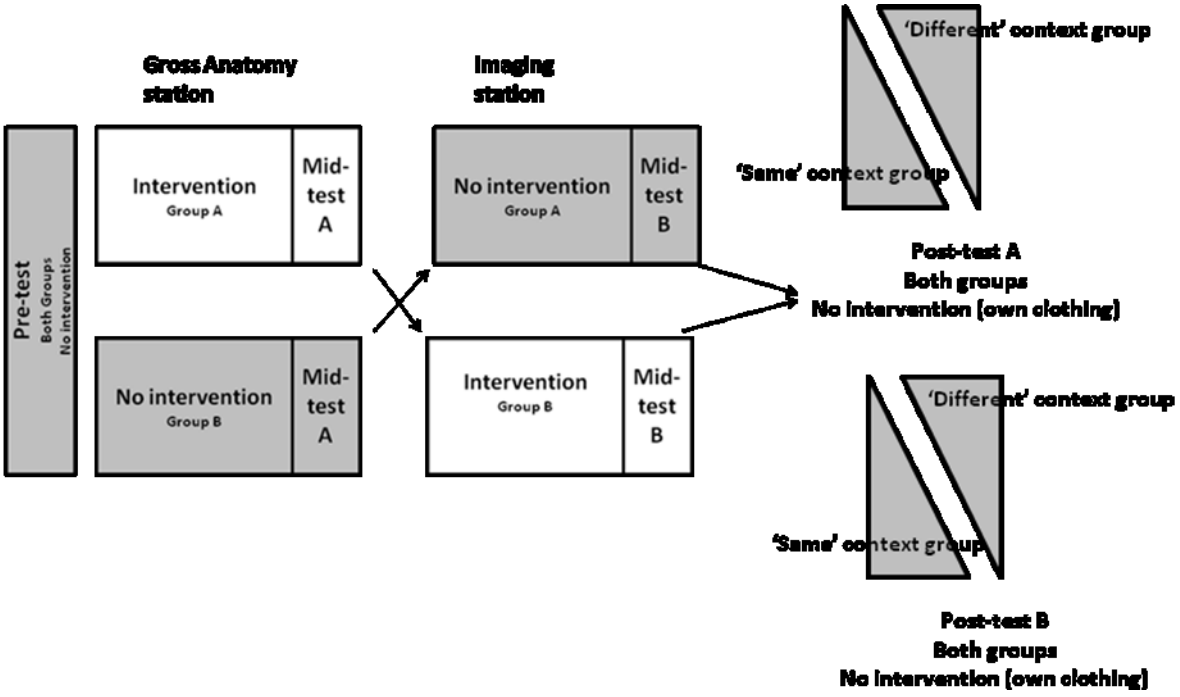

Figure 8: Cross-over design.

This diagram illustrates the teaching session and the allocation of 'context' (clothing) for each component.

= students wearing scrubs

= students wearing own clothes

All students sat a pre-test on the gross anatomy and imaging of the kidney. This was used to establish a base-line of their existing knowledge of renal anatomy and to check randomisation of anatomy groups (described later). Students were likely to have a reasonable amount of prior knowledge as the renal system is present on most pre-university biology syllabii. The pre-test also provided information on the equivalence of the groups in terms of academic performance. The pre-test was administered 2 weeks before the teaching session on renal anatomy. This was so that the students had not studied the

kidney immediately prior to the teaching as they often do. The pre-test was computer based. Students completed the pre-test in their own clothes.

The pre-test consisted of 12 multiple-choice questions (MCQs) based on 3 images, i.e. 4 questions per image. All tests were administered online, using the university's Virtual Learning Environment (VLE).

Students were divided into the eight anatomy teaching groups to which they had been randomly assigned at the start of the academic year. All consenting students were issued with ceil coloured scrubs. This colour was chosen as it matched that which the students will wear in the local hospital environment, thus making the simulation more authentic.

Students wore either scrubs or their own clothes for the gross anatomy teaching session, depending on the group allocation. A gross anatomy specific mid-test was then completed immediately after the teaching session. This test followed the same format as the pre-test, 12 MCQ questions based upon 3 images.

Students then changed into the opposite clothing and attended a self-directed session on the imaging of the kidney. An imaging specific mid-test was then completed, following the format described previously.

Five weeks after teaching, students completed two post-tests, one gross anatomy and one imaging specific. All tests were standardised and conducted online, thus giving students an instant score. Each test contained 12 multiple-choice questions, each based upon a photograph of a prosection, or a radiological image.

Students completed the post-tests in their own clothes. Whilst it would have been ideal to have tested students in both their own clothes and scrubs, this proved to be impractical in response to student feedback. Students raised concerns over the provision of changing facilities and the time that would be needed to change.

5.3.3 Data collection and analysis

As tests were created using Blackboard™ software on the VLE, data were collected automatically. Test scores were downloaded from the VLE as Microsoft Excel files. Data analysis was performed using a paired t-test (SPSS v15.0 for Windows).

5.4 Results

Of the 93 first year medical students, 82 completed all aspects of the study.

Reasons for non-inclusion were:

i) Absence from the pre-test (1)

ii) Absence from the teaching session & mid-test (3)

iii) Absence from the post-test (3)

iv) Failure to submit responses on on-line tests, thus registering the test as 'still in progress' (2)

v) Unwillingness to wear scrubs (2)

There was no statistically significant difference between groups' pre- test, or mid-test scores. Table 9 shows the percentage scores for both groups' pre-, mid- and post-tests.

Test	Group A	Group B
Pre-test	33.2	39.6
Mid-test A (Gross anatomy)	90.3	85.4
Mid-test B (Gross anatomy)	83.7	82.4
Post-test A (Gross anatomy)	73.9	75.9
Post-test B (Gross anatomy)	80.5	77.4

Table 9: Percentage scores for pre-, mid- and post tests for the two groups (A & B).

The post-tests are highlighted to show whether in which context the tests were taken. This could be either 'same' or 'different' context to that in which the content was taught.

Tested in same context = 

Tested in different context = 

Post-test scores for students' completing in the 'same' and 'different' contexts as that in which they learnt were aggregated. Students completing post-tests in the same context had knowledge decay of 4.2%. Students in a different context had a 10.1% decay of knowledge (see table 9).

Context	Percentage decline in recall from mid-test (%)
'Same'	-4.2%
'Different'	-10.1%

Table 5.2: Percentage difference

This table shows the percentage difference in post-test scores between the two groups. The percentage shows the decline in test scores from mid- to post-test.

Scores for groups tested in the 'same' and 'different' clothing contexts to that in which learnt they were aggregated. The results show a significant difference between the 'same' and 'different' groups, by paired t-test; p=0.04

5.5 Discussion

Students examined in the same context as that in which they were taught recalled significantly more information ($p=0.04$). The results of this study demonstrate that recall is context dependent to some extent. The effect size of 0.27 is small, and undoubtedly there are other factors affecting recall. However, for an educational experiment this is a significant result, especially given the seemingly trivial context of clothing. The effect size is also of significance when the complexity of educational and learning processes are considered. The recall was significantly higher than the level achieved by Godden and Baddeley[205]. This is likely to be a result of using content relevant to the participants, as well as an authentic setting. As mentioned previously, a significant drawback of the Godden and Baddeley[205] study was the non-real world setting, whereas in this study the context simulated is authentic.

These results imply that if recall is better in the context in which it is learnt, then medical education should move towards clinically simulated contexts in the pre-clinical years which have relevance to future clinical practice, even where the context appears irrelevant to the students' current learning. Learning has been described, particularly in the workplace context, as being a result of purpose and direction [213]. The students in this study were recalling information that they are required to learn in order to pass future examinations and certainly use in future clinical practice.

This effect size could have been higher had the study design been optimised. The reason for the slightly sub-optimal design in practice is worth noting as it is an issue for everyone implementing medical education research. Students completed the post-test approximately 5 weeks after the teaching session. As the post-tests were only short MCQs the students were not willing to change into scrubs for such a short period of time. This would have been less of an issue if the groups had not needed to change between tests, so that some students completed tests in the same context to that in which they learnt, and others in the same. But the research was carried out in a practical, real world setting, where student views and responses are both real and important. This indicates that a perfect study design is not always practicable, especially given the intensive nature of medical courses. I acknowledge that the study design was imperfect, and that this may have impacted on the effect size: however, study design must not disadvantage students, even if only in their own perception.

It is interesting and relevant to consider the impact of testing itself. It is said that in the long term large group teaching is not effective in terms of student learning[214]. Attention usually wanes after about 20 minutes and it is thought that often only 20% of the information taught can be recalled later. Knowledge retention following conventional teaching often decays at an undesirable rate [174]. Educators are encouraged to establish what learners already know is encouraged as it allows the content delivered to be better matched with the learners' needs. This suggestion may be of relevance when utilising a pre-test in medical education, and in particular in research projects. Although testing may not be popular with students, it provides an instant feedback opportunity which might not be available ordinarily. Post-teaching testing is a valuable way to improve retention of knowledge[174-176]., as it has been shown to slow decay of knowledge [175]. Karpicke and Roediger found that it is

the “repeated retrieval induced through testing, and not repeated encoding during additional study, that produced large positive effects on long-term retention.” Cognitive psychologists have named this phenomenon the ‘testing effect’. The testing effect shows that testing directly affects learning by promoting better retention of knowledge [176].

These results suggest that simulation approaches in anatomy teaching have real value. This is in line with suggestions that a variety of learning strategies should be utilised in order to promote deep learning [215] . This seemingly irrelevant detail of clothing when simulating a more clinical context demonstrates that for a high degree of authenticity in simulation approaches, small details are important. Clinical simulation has been reported to be on the point of having a significant impact on medical education [178], and is thought to provide a safe, supportive educational environment [216]. There is value in re-creating some aspects of the real clinical situation [178], whether it be mimicking the clothing aspect of the clinical context, re-creating aspects of the built environment such as simulated wards, or using expensive, high-fidelity simulators.

It has been suggested that there has been a decline in the standards of anatomical knowledge among medical graduates[72, 185, 217]. Students may declare an inability to recall anatomical information, taught in the classroom setting, when challenged in a clinical setting. For many students it is easier to deny being taught than to admit not remembering, or to propose a partial hypothesis. This is may be in part due to a fear of the consequences of being publicly shown to be incorrect. But it may also be a consequence of the practice setting being different from the previous learning setting. Simulation allows the novice and

the expert to practice and develop skills with the knowledge that mistakes carry no penalties [178], and may aid subsequent recall.

The results of this study may represent a reference to the notion of the 'legitimate peripheral participator' as described by Lave and Wegner [218]. Lave and Wegner describe learning as a process which is social in nature. Legitimate peripheral participation refers to how newcomers to a subject become integrated into a community of practice. During the learning process the newcomer moves from peripheral participation to eventual full participation, the success of which is dependent upon both social dynamics and the power structure. As the newcomer moves from the periphery of this community to its centre, they become more active and engaged within the culture. This will continue until they eventually assume the role of an expert. Learning occurs *in situ*, which may suggest that all learning is in fact contextual and embedded within the social and physical environment of the community in question.

Lave and Wegner propose that situated learning ' is not an education form, much less a pedagogical strategy or a teaching technique' [218]. The authors argue that learning as it normally occurs is a function of the activity, culture and content in which it occurs (the community). In other words, it is situated. This is in contrast with classroom learning activities which for the most part involve knowledge which is abstract and out of context. It appears that the situated learning process is unintentional, rather than deliberate. The literature is perhaps implying that knowledge is situated in the practices of the community, rather than something which exists in books. Eraut describes this as "learning from experience" [213], and attributes learning to the social interactions within the place of work.

While it would be ideal for medical students to do more 'situated learning' within the community of practice, this is hard to achieve for both practical and pedagogic reasons. Students may need a basic science background on which to build their clinical knowledge so that they can eventually become legitimate members of the clinical community. Simulation may provide a way to bridge the gap between classroom learning and legitimate peripheral participation. It may not be possible to replicate the social aspects of clinical community, but environmental and contextual factors are more easily reproduced, and have a measurable effect on recall as this study has shown.

Using clothing as a method of simulating a more clinical environment may be a cost effective mechanism by which simulation can be introduced into the undergraduate curriculum, particularly in the non-clinical university environment. Simulation in high-fidelity purpose built suites can be expensive, and expert clinical facilitators would require payment. The approach described in this study enables simulation to be cost-effective while still having an educational impact.

The take-home messages from this study are:

- Recall is context dependent
- Seemingly trivial contexts, such as clothing, are relevant in simulation approaches
- An optimal study design is not always possible in real educational settings

5.5.1 Limitations and future work

As previously described, the change to the design of this study in response to student feedback was flawed. However, this research was conducted in a real world setting. A comparative study is now underway to further investigate the use of context in the medical curriculum.

Part 2: Students' views on wearing scrubs

5.6 Background

Part two of this study describes the attitudes medical students have towards professional clothing, particularly personal protective clothing such as scrubs and lab coats. The impact such clothing has on education and learning in context is also addressed.

5.6.1 Conceptual framework

My observation of the intervention revealed that there was a significant level of excitement amongst the students when wearing scrubs. Students had posed wearing their scrubs and taken photographs, which were uploaded onto social networking sites. In light of this, I decided to conduct focus groups to investigate this.

5.7 Methods

5.7.1 Ethics

Ethical permission for part two of this study was granted by the School of Medicine and Health's Ethics Sub-Committee.

5.7.2 Recruitment and consent

All first year medical students were invited to participate in a qualitative research project. Recruitment and consent procedures were conducted as described in Chapter 2. Students who participated or declined to participate were asked to attend the focus groups.

5.7.3 Data collection and analysis

Focus groups were conducted by the author (GF), utilising a semi-structured approach. Four focus groups were conducted, each lasting approximately 40 minutes. Focus groups followed the qualitative data collection procedure outlined in Chapter 2. Transcriptions were produced by the author (GF) and an administrator (HT). I analysed qualitative data using grounded theory, as described in Chapter 2.

5.8 Results

Twenty seven students (14 females, 13 males) participated in part two of this study. Two principal themes emerged from the qualitative data; ‘clothing’ and ‘educational impact’.

Within the ‘clothing’ theme there were 5 subthemes, these were ‘ symbolism’, ‘enjoyment’, ‘scrubs versus white coats’, ‘acting up’ and ‘self-consciousness’. Within ‘symbolism’ the

sub-theme of ‘professional trademarks’ emerged. There were 3 subthemes emerging from the ‘educational impact’ theme; these were ‘memory and recall’, ‘impact of dissection’ and ‘instant feedback’. There was a further mini- subtheme within ‘memory and recall’; this was ‘novelty factor’.

Clothing

Students discussed the issues which arose when wearing scrubs as part of the intervention.

Self-consciousness

Some students reported feeling self-conscious whilst wearing scrubs.

“I was concerned that...they wouldn’t fit. It sounds odd but I couldn’t help it, I find it rather difficult to get a lab coat to fit me.....”

“I thought I looked awful in it. I felt like a man.”

Enjoyment

Students appeared amenable to the idea of wearing scrubs, and stated enjoyment.

“To be honest with you I quite enjoyed it.”

“I thought they were quite good; I quite liked the idea of wearing scrubs.”

“I feel excited about wearing them and I enjoyed it.”

Symbolism

Wearing scrubs revealed the symbolism clothing has within the formation of professional identities.

"I don't know, they make it seem more real don't they? The scrubs bring you up to date because now doctors in hospitals wear scrubs don't they?"

"Just feeling that you kind of moving from being just a student to someone who's potentially going to be a doctor. I'm thinking 'right we're wearing the proper garb of a doctor'...I really enjoyed it and have to admit I did take a few photographs of myself in it... because it's the first time like I say I'm wearing like the professional garb of the profession I'll be going into so that's why I quite enjoyed it."

"[scrubs are]... more clinical I suppose...

"... [I got a] little excited. Because we're wearing something different, and we've got to sort of dress up like proper doctors and dress like TV characters."

Professional trademarks

Whilst describing the symbolism of clothing, students made references to white coats being out dated. Comments referred to the white coat being associated to either the laboratory

environment or scientists. Students directly compared this to how scrubs are now associated with medics as opposed to the white coat.

“Although lab coats are quite nice in this...1950s scientist kind of way...”

“... the white coat is kind a thing of the past in the hospitals and it’s purely in the laboratory that we wear them...”

Scrubs versus white coats

For any laboratory based teaching students must wear a white lab coat. Students fed back their perceived differences between wearing the scrubs and the lab coat. Comments from the students showed a strong preference for wearing scrubs instead of lab coats.

“I saw it as quite different, well similar but different in the aspect that now we’ve been told that the white coat is kind a thing of the past in the hospitals and it’s purely in the laboratory that we wear them so I did actually see it as different, I did have a kind of different feel to it and it also again obviously I’ve worn the white coat quite a bit now so obviously I suppose they’re kind of, a bit more familiar with it, the kind of newness has kind of worn off.”

“Yeah I think they’re [scrubs] better than the lab coats.”

“I suppose if you wore them all the time it might not have that same buzz about it, saying that I’ve only got one lab coat because they didn’t have another one that fitted me...”

“... I found them nice...I moved in them more easily than a lab coat ...”

“...we had this kind of buzz about it because we were all there in scrubs and you know we were all buzzed up by it. I don’t know, I know a few of us, were like ‘damn we have to give them back’ because we would like to have kept them...”

Acting up

Students commented that whilst wearing scrubs they felt like a ‘proper doctor’. Others regarded wearing scrubs as pretence.

“[Wearing scrubs]...allows you to depict yourself as what you aim to be in the future and it allows it to happen instantaneously.”

“...it put you in the mindset of being a doctor, being able to achieve being a doctor at the end of the day.”

“When you are a student you’re like ‘I know nothing’ then you put on a pair of scrubs and then you’re like ‘one day I will know something’.”

“I felt a bit pretentious...”

“I felt like I’m pretending to be something I am not”

Educational Impact

Students discussed the educational impact that they perceived the scrubs intervention to have had.

Memory and recall

Students described how they perceived the intervention to have had a positive effect on recall through the use of contextual learning and testing:

“I think it helped me remember it quite a lot...”

“I can remember the whole day and the things that we learnt...”

“I don’t know if, how it affected my memory recall but maybe I probably put more...concentration into answering the questions.”

Novelty factor

Some students regarded any impact that the use of scrubs had as being down to the novelty factor and it being ‘something different’:

"...we could have been wearing clown suits and I think we still would have remembered it because it was different..."

"I do remember all the stuff from that day but....it could just [be] because it's something different."

5.9 Discussion

Although clothing may appear to be a trivial consideration in an undergraduate setting, the comments made by students in this study indicate otherwise. Students place significant weighting on clothing as it appears to define the professional identify of those they aspire to be, namely practising doctors.

Clothing has long been recognised as a communicator of information during social interaction and as an aid in the establishment of self-identity[219], a role acknowledged within society. For the students in this study, most of whom were young adults, they were in a formative stage of their identity development. Daters (1990) states that, as adolescents search for a self-image and identity, they may become preoccupied with experimentation in their appearance. Clothing importance and clothing interest have been recognised to be at their peak during adolescence[219]. Furthermore, physical, social, and psychological factors may influence how one sees, values, and uses clothing [219]. These comments may shed some light on the emphasis students place on wearing scrubs. However, clothing has been linked to social acceptance[220].

As I described in part one of this study, students are striving to become legitimate participants in their community of practice. Acceptance into this community, by both its members, medical staff, and its service users, the patients, will undoubtedly by an important consideration for them. Those that students perceive to be masters within their community of practice wear scrubs, thus by mimicking this uniform, students may feel engaged and accepted within the community.

Students' comments regarding how wearing scrubs made them feel like a 'proper doctor' may provide evidence of professional identity formation.

Within the data presented, students' association between scrubs and the medical profession was distinct, considering that the abolition of white coats from the wards was a recent event. References to 'TV doctors' were made, such characters and their dress may be partially responsible for the strong association made by students. An example of this might be 'Scrubs', the 2001- 2010 cult comedy drama, which is popular with medical students. Within this TV series, doctors wear the ceil scrubs, as utilised in this study, and allied health professionals wear other colours including pink.

Feinberg and colleagues[221] believe individuals to use clothing to define and communicate their social identities to others. They report that observers should be able to read and show consensus about the social information present in clothing cues, and there should be substantial agreement between the perceived meaning of the cues to observers and the actual social identities of the individuals but only if the clothes are specifically selected by

the individuals to represent themselves[221]. Students appear to be adopting this strategy to communicate their future professional identities, perhaps best evidenced by their posting of photographs of themselves dressed in scrubs being placed on social networking sites.

Clothing is thought to be a significant social symbol used by individuals in identity definition because (1) clothing is used on a daily basis, (2) clothes constitute a frequent public display, and (3) clothing choice is an easily manipulated symbol[221]. Students view professional clothing as a status symbol. Individuals form impressions of others based on clothing cues and observers behave differently toward people depending on the clothing they see them wear[222]. I believe that the students will be aware of this, as it is the culture into which they have been socialised. Additionally, individuals appear to prefer clothing which communicates images similar to the images they have of themselves[222]. The participating students wish to view themselves as doctors, doctors wear scrubs and hence there is a ‘buzz’ about being given the opportunity to dress as the professional they wish to be.

Status and group cohesion is important to medical students. For example, medical students at Durham University, which is a collegiate university, are often criticised by the staff at their colleges for not engaging in college activities. Medical students have their own societies (MedSoc) and sports teams. This is perhaps best highlighted by the data where students report a preference for scrubs because white coats are seen as the professional trademark of scientists.

With respect to the educational impact of the intervention, comments were more concentrated. During the focus groups students were less interested in discussing the educational aspects of the study, but were more focused upon describing the dressing-up process and how they were able to act like doctors. This anecdotal evidence offers an insight as it demonstrates to educators how clothing is an important aspect of identity formation and simulation of the clinical context.

Students suggest that the educational benefit from this intervention was not as a consequence of wearing scrubs, but because the session was ‘different’, thus it was as a result of the novelty factor. This may well be true. However, their strong association between the clinical environment and scrubs indicates that the novelty factor might not be the whole case. For the impact of the scrubs to be sustained over the prolonged period of this study would be unusual. Saying this, it might be possible that a ‘learning landmark’, described in Chapter 4, had been created and thus led to the retention of knowledge.

I have previously described[90] ‘‘learning landmarks’’: vivid experiences which are memorable in themselves, and which then provide access to the educational content associated with that context. The simulation of a clinical environment within undergraduate anatomy teaching and the dressing up as a ‘proper doctor’ combined with fresh animal tissue dissection could all individually, and certainly collectively create the opportunity for a student to forge a learning landmark. As described above, this vivid experience sticks in the memory of the student and appears to enable retrieval of information from the learning context in which the memory was forged. Students’ comments that they remembered the content taught because it was something different are, in my experience, indicative of the

formation of a learning landmark. This study may have cemented their future professional identity as they were able to act like a doctor, and this positive memory will further reinforce the learning landmark.

I cannot end this discussion without considering the practical implications of such an intervention; that is, issues impacting upon willingness to participate. Some students were self-conscious about wearing scrubs. Although scrubs are compulsory for Phase 2 medical training (years 3-5) as they are the uniform for medical students, it would not be a positive learning environment if students were forced to wear scrubs as part of a study. The unwilling students were a minority, but this serves to highlight that an opt out or alternative should be offered to students. Forcing any student to dress in a manner which causes them distress will not promote learning, and will go against the creation of positive learning landmarks. It may even go as far as to create a negative association. “Identifying and understanding self-esteem builders such as clothing comfort, may help...educators find specific methods to inculcate constructive behaviours[219].”

This study has highlighted how clothing can contextualise information; such information could be the anatomy as taught in this study, or information about the role of a person within a given environment. For students, professional dress appears to define environments and the students’ identities. Clothes provide cues toward which individuals can make indications allowing them to negotiate their identities. Futhermore, ‘clothing enables them to understand one another’s identities in a given situation’[220]. This identity negotiation allows for the subsequent interaction to proceed more smoothly. Appearance may either impede or facilitate social interaction as individuals make indications toward one

another's appearances. Future interactions, with patients and other staff, are a consideration for students, even at this early stage in their career. All the academic study undertaken is based upon the goal of becoming a doctor and the students' professional development is emphasised by faculty from entry to medical school. Professional development teaching incorporates instruction on standards of professional dress. Contemplating this enables educators to comprehend the students' reasons for regarding professional clothing so highly.

5.9.1 Limitations and future work

The sample size for this qualitative aspect of the study is relatively small, however opinions appeared to be consistent. Further work is planned to consider clothing and professional identity formation.

Chapter 6: The Conscientiousness Index

6.1 Background

This multi-phase study assessed the use of the Conscientiousness Index in undergraduate medical education. This study is in 5 parts, and will be described in Chapters 6 and 7.

6.1.1 Contributions to this Chapter

Professor John McLachlan devised the Conscientiousness Index. He analysed the data for part one of this study. Professor Jane Macnaughton contributed to the manuscript published in Academic Medicine[93]. I was responsible for the literature review, data collection and revising the manuscript for part one.

I devised the study for part two, collected and analysed the data, and wrote the manuscript published in Medical Education[92]. Dr Marina Sawdon contributed to the data analysis and edited the manuscript for part two. Dr Laura Clipsham contributed to the data collection and reviewed the manuscript for part two of this study. Professor John McLachlan also reviewed the manuscript for part two.

In addition to the two papers published there have been presentations and workshops at International conferences[92, 135, 223-229].

Part 1: The Conscientiousness Index: A Novel Tool to Explore Students' Professionalism

Measuring professional behaviour is problematic not least because the concept of professionalism is difficult to define. Part one of this study describes a measurement tool that does not rely on qualitative judgements from respondents but nonetheless clearly correlates with individuals' subjective views about what constitutes professional behaviour.

6.2 Introduction

There is evidence that negative student behaviour during undergraduate programmes is related to the likelihood of subsequent negative behaviour in later careers or post-qualification practice[1 - 5]. This, in addition to recent concerns about public perceptions of the medical profession[6], is the main reason why the issue of professionalism – how to teach, assess and evaluate it – has become such a hot topic in medical education. Research has focused on attempts to define professionalism, [8] for without a definition or clear understanding of what professional behaviour consists (it is argued), it cannot be taught and certainly cannot be measured. Professionalism is a slippery concept and definitions are difficult. Most comprise a list of attributes variably including altruism, honour, integrity, excellence, accountability, respect for others, [9] compassion, continuous improvement, and working in partnership[10]. Each of the attributes requires further explanation and agreement about what it means in the context of undergraduate education for it to be of any use either to guide teaching content or measurement[11]. These complexities may lead to differing interpretations of professionalism on occasion.[12]

While there is not one single definition that is universally accepted, staff and students in our medical school share a working understanding of professionalism based on the General Medical Council (GMC) publication '*Good Medical Practice'* [13]. This document is given to all students on induction, and course outcomes relating to professionalism are drawn from it, explicitly and often verbatim. Students are advised that the principles in '*Good Medical Practice' [230]* apply to them, and breaches could represent grounds for disciplinary proceedings.

As Durham University utilises this document[230] as the source of its working understanding of professionalism for the purposes of this study, an overview of the document is in order. It is described as guidance rather than as a statutory code, although sentences containing the words 'you must' are to be regarded as "an overriding duty or principle." Six key principles are identified as the essential summary. The first of these is making the care of a patient doctor's first concern. The second is the protection and promotion of the health of patients and the public. The third is the provision of a good standard of practice and care, which includes keeping knowledge and skills up to date, working within personal limitations, and working co-operatively. The fourth is treating patients as individuals and respecting their dignity, which includes treating them considerately, but also with confidentiality. The fifth deals with working in partnership with patients, listening to their concerns, giving them information appropriately, respecting their decisions and supporting their self care. Finally, honesty, openness and integrity are required, and under this heading are included non-discrimination, meriting trust, and acting without delay if the doctor or a colleague is putting patients at risk.

Altruism is not specifically named as an overriding duty, although 'making the care of the patient the doctor's first concern' does carry this implication. Further detailed guidance is provided on providing good clinical care; maintaining good medical practice; teaching, training, appraising and assessing; relationships with patients; working with colleagues; probity and personal health and self care. As this indicates, the document includes well-accepted criteria for professionalism, although it does not define the term *per se*.

Since professionalism is a qualitative entity, measurement in a scalar fashion that requires assessment of one person's degree of professionalism against another's is problematic.[14] As a result, attempts to construct meaningful scales for the measurement of professionalism might seem doomed to failure.[15] In addition, if professionalism is thought of as a competence[16], like examining the cardiovascular system or like some aspects of the teaching of communication skills, it then becomes the subject of a training approach in which students might be encouraged merely to *practice* honest behaviour or good partnership. Assessment, accordingly, becomes focused on observing what has been practised, and medical students are good at regurgitating required behaviour as well as facts. Thus we may not be sure that what is measured as a result of a competency-based approach is what the student might actually do in practice[17].

In the light of these challenges, a scale was constructed for an aspect of professional behaviour that has two advantages in this problematic context: first, it is amenable to meaningful quantitative analysis; and second, it records spontaneous behaviour. This scale has been called the Conscientiousness Index (CI) of student performance and was constructed during a full academic year (2006-07) for medical students in Year 1 and Year

2. (The word “conscientiousness” was used because conscientious students will do those things that are required to score well, while less conscientious students will not.) Diligence is another possible synonym for conscientiousness. As will be clear from the methodology, no qualitative judgements were required for the collection of CI data, so each individual student was legitimately given a score for conscientiousness that could also be compared directly with the scores of other students.

Departmental interest in Conscientiousness stems from the well known psychologist phenomenon ‘The Big Five’. The Big Five are personality traits, thought to be stable after age eighteen. These factors are: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism.

Due to their stability after age 18, any relationship between such traits and professionalism is of interest due to the age of medical students. Furthermore, as these five factors have been found to contain and include most known personality traits and are assumed to represent the basic structure behind all personality traits.

6.3 Method

6.3.1 Ethics

This study was not designed as an interventional research study: all the data involved in the CI were already being collected routinely, and the only difference to our standard practice was to gather it together in one place.

In the School of Medicine and Health, educational studies that do not require unusual approaches and that do not compromise the anonymity of the students have been considered as not requiring prior ethical approval from internal review boards. [18] However, the applicability of that general principle to this particular study was confirmed in writing by the Chair of the School Ethics Committee.

6.3.2 Recruitment

Students were advised orally and through the virtual learning environment (VLE) at the beginning of the year that their teachers would be scoring them using the CI, and that the CI had no summative or formative role. They were also informed that their anonymity would be preserved in all analyses of the data, and that no academic harm would come to them from this process.

In the year under consideration (2006-7), there were 106 first year students and 118 second year students.

6.3.3 Data collection

Faculty members or administrative staff in each course taken by students recorded the data which led to the award or deduction of specified numbers of *conscientiousness points* for each student for the reasons summarised below.

Awarding or deducting conscientiousness points

Attendance

Attendance at compulsory teaching sessions during Year 1 and Year 2 was recorded by teaching faculty or administrative staff through the use of registers. At the start of the school year, each student was awarded 50 points for each of the three terms that year. One point was then deducted from this total for each unauthorised absence. (The school authorises absences in advance or in retrospect for good cause, such as hospital appointments or illnesses).

Submission of data

Students were advised in writing on two separate occasions with a deadline of the necessity to provide specified information on their Criminal Records Bureau status (i.e., previous criminal convictions) and immune status (in the U.K., students are required to declare on entry if they are positive for a defined list of transmissible diseases, including HIV/AIDS and hepatitis B). Students were awarded two points for submitting the required information before the deadline, one point for submitting partial documentation before the deadline, and no points for failing to submit any of the required information before the deadline.

Evaluative feedback

Student feedback was gathered using the university VLE, described earlier. Students keep track of timetables, access teaching materials, complete both formative and summative assignments and submit course evaluation feedback through the system. The VLE has the facility to track those students who have submitted course evaluation feedback without identifying the content of their feedback. Students were advised that failure to submit feedback would be taken into consideration in the monitoring of student conscientiousness. Students gained one point for each evaluation they submitted, but did not lose points if they failed to submit.

Summative assessments

Students were trained in the use of the Speedwell Learning Systems Optical Mark Reader (OMR) format employed in the medical school. The OMR enables students to mark responses to multiple-choice and extended-matching-item-format assessments (given in connection with their classes) on pre-printed sheets, using a pencil. An introductory training session, with practice, and subsequent individual detailed feedback on incorrect usage, was provided. The training session information was made available through the VLE. A full formative examination using the OMR was conducted, again with individualised feedback. Students were then awarded one conscientiousness point for (1) each correct use of anonymous examination numbers during summative exams and (2) each correct completion of optical mark reader (OMR) responses during summative exams. All assessments that involved the OMR and that contributed to progression decisions were included. There were two such assessments in January and two in May, which is the usual pattern for such assessments.

Summative assignments

Subsequent to a training session on how to submit assignments, the posting of assignment submission guidance on the VLE, and a practice formative opportunity, students received one conscientiousness point for each assignment correctly submitted on time. Those who failed to submit an assignment on time, without prospective or retrospective authorisation for good cause, received no points. All assignments that contributed to progression decisions were included. There are three of these in each academic year.

Voluntary participations

Those students who participated in voluntary activities relating to the medical school were awarded one point for each separate activity. Such activities included sessions devoted to widening the range of candidates who apply to medical school and sessions included working with school students on visits to the medical school, assisting at extended “master classes” in which schools students from deprived areas attend events over several days and given an introduction to the kind of activities found in undergraduate medical courses, and visits to local schools to encourage interest and answer questions.

Uncategorised events

In addition to the above categories, students could receive and lose points for individual, uncategorised events, as advised by the Programme Manager (a senior Faculty member who is responsible for the day to day administration of the programme). These events were generally unique to a student. Positive events included responding professionally to a genuine medical emergency, advising staff of a possible examination impropriety, and

advising staff of a possible breach of patient confidentiality. Negative events included reading but failing to respond to repeated e-mails from staff, and attending a teaching session in an unfit state. These negative events would not be significant enough in themselves to trigger a critical incident report (see below) or a fitness to practice procedure.

6.3.4 Validity

6.3.4.1 Establishing the validity of our approach

The GMC has listed a number of the above categories in their document *Medical Students: Professional Behaviour and Fitness to Practice*[19], supporting the face validity of our use of them as indicators of undergraduate professional behaviour. However, in order to establish the concurrent validity of the approach (whether the CI correlated with other views or estimations of professionalism) any relationship between the CI and staff views on individual students' professional behaviour and with data on critical incident reporting, were explored.

6.3.4.2 Validity: Correspondence with staff judgements on professionalism

The top ten, middle ten and bottom ten students in each year, on the basis of their CI scores, were classified. The RAND function in Excel (Excel 2003 for Windows) was used to randomise the order of their names. A group of nine experienced staff members was asked to express an expert judgement on the professionalism of these students in each of the two years, using the three options listed in the following paragraph. These staff members were isolated from knowledge of the CI scores of students throughout the year. Staff members

who made these judgements were unaware of scores as they were not responsible for collating the CI data. Furthermore, when they made expert judgements they received only the names of the students on a questionnaire, no scores were attached. Since our cohort of students is relatively small (of the order of 100 - 120 per year), and their two years with us are spent on a residential campus that is also the base for our staff, there is close and frequent contact between staff and students. The inclusion criteria for staff were (1) experience in working with medical students in general, (2) familiarity with the GMC definition of professionalism, and (3) close and repeated contact with students in teaching and support capacities throughout the relevant time period. The nine staff members included eight members of the academic staff, at lecturer level or above. The ninth was a senior technician who also played a teaching role. The academic staff included teachers in classes entitled Personal and Professional Development, Medicine in the Community, Anatomy, Physiology, and Clinical Skills. They received no additional compensation for the task and received written guidance as to the general purpose of the exercise and a summary of the outcomes after the completion of the draft manuscript.

The staff were asked to choose one of the following options:

- I am happy with the professionalism shown by this student.
- I have some concerns with the professionalism shown by this student.
- I do not know this student well enough to comment.

The staff choices for each student were then compared. To further explore the continuous nature of the relationship, a professionalism index (PI) was calculated for each student, consisting of the “happy” score minus the “concerns” score. For instance, a student

receiving five positive evaluations, three negative evaluations, and one “don’t know” would have a PI of +2. The maximum score is 9, since 9 evaluators were involved. No adjustment was made on the basis of the absolute number of evaluations made: in other words, a student receiving 5 positives and 4 negative evaluations was given a score of +1, as was a student receiving 2 positive evaluations, 1 negative valuation and 6 “don’t knows.” The correlation coefficient between the CI and PI scores was then calculated.

6.3.4.3 Validity: Correspondence with critical incident reports

A critical incident reporting system is in place in the undergraduate medical programme, modelled on the UK National Learning and Reporting System which was created by the National Patient Safety Agency. A critical incident report is completed by staff members or students when they observe and choose to report a critical incident. Students are then invited to respond and reflect, using forms adopted from the National Health Service reporting forms. The occurrence of such recorded critical incidents for each student was also compared with his or her CI score.

6.3.5 Reliability

To explore the reliability of the results, the academic year was split into two, and performance on the first half year was compared to the performance on the second half year, for each of the two years under study. The Spearman Rank Correlation Coefficient was calculated for each group.

6.4 Results

6.4.1 Distribution of the Conscientiousness Index scores

Conscientiousness scores ranged from 153-205 for Year 1 students and 116-195 for Year 2 students. These scores were converted to percentages of the maximum possible scores and are displayed as histograms in Figure 9, Figure 10 and Figure 11. Even casual comparison of these figures reveal that the percentages are quite similar. The descriptive statistics are shown in tables 10-12.

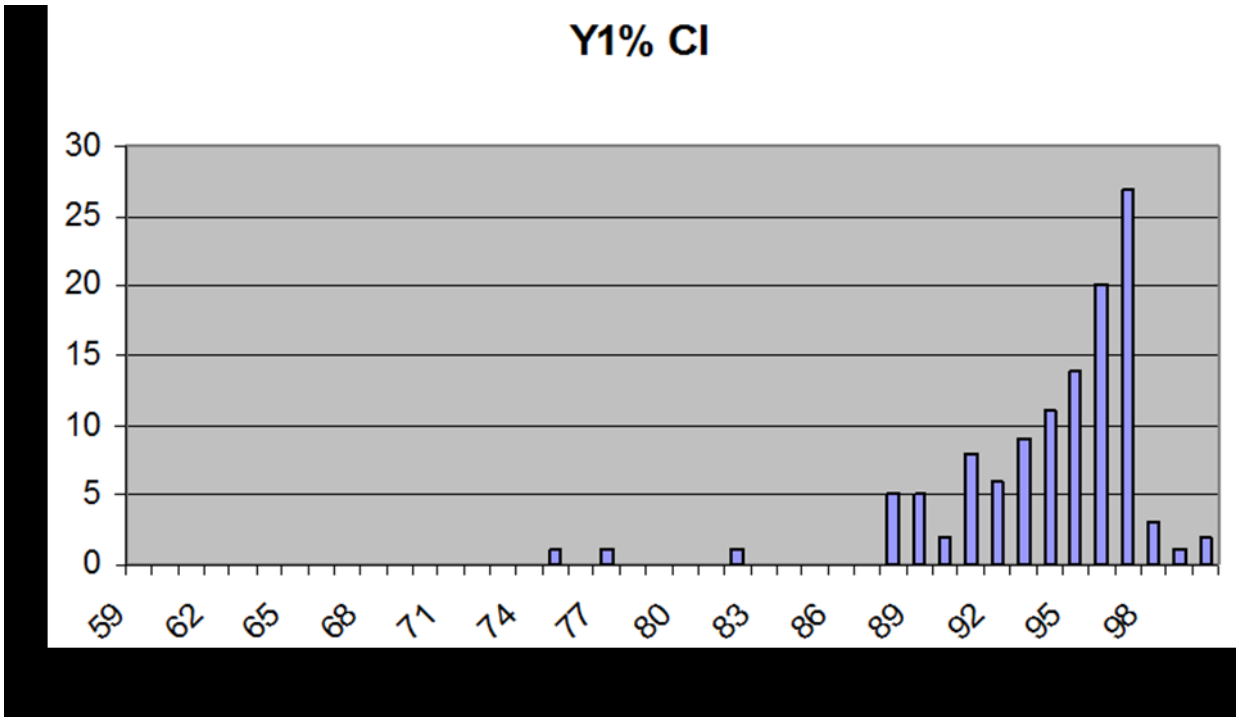

Figure 9: The Conscientiousness Index scores shown as percentages of the maximum possible score, for Year 1 students, 2006-07.

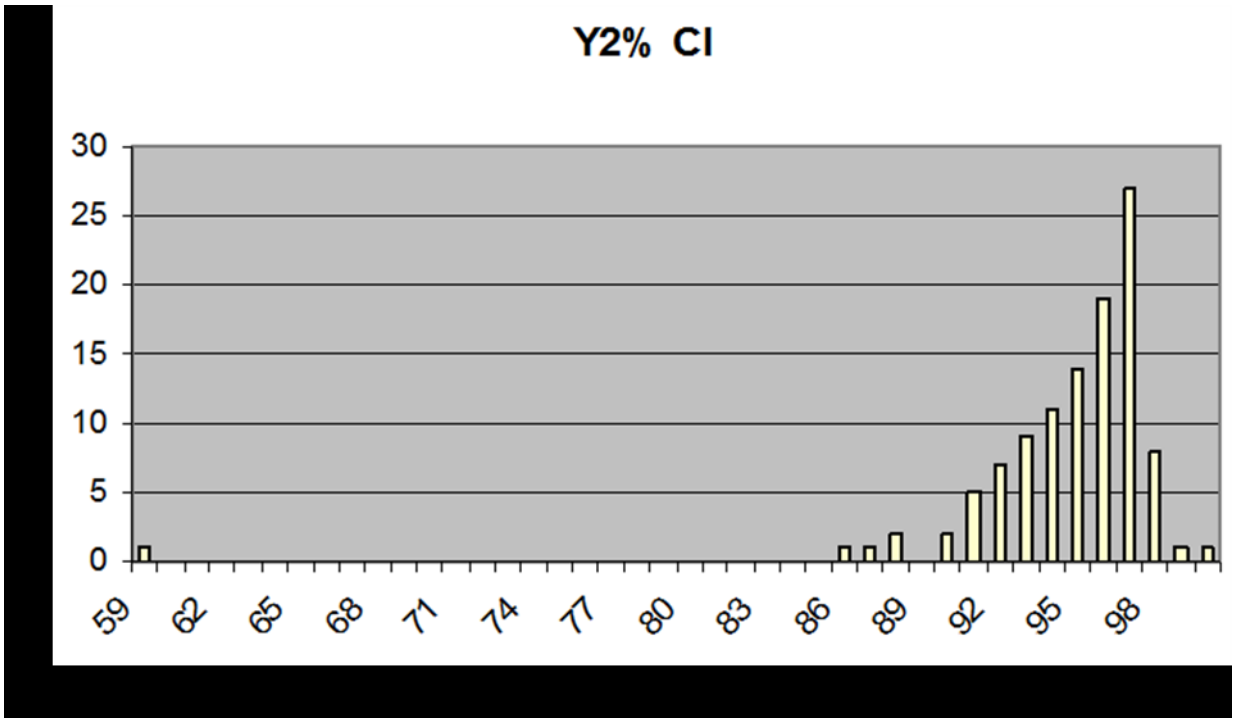

Figure 10: The Conscientiousness Index scores shown as percentages of the maximum possible score, for Year 2 students, 2006-07.

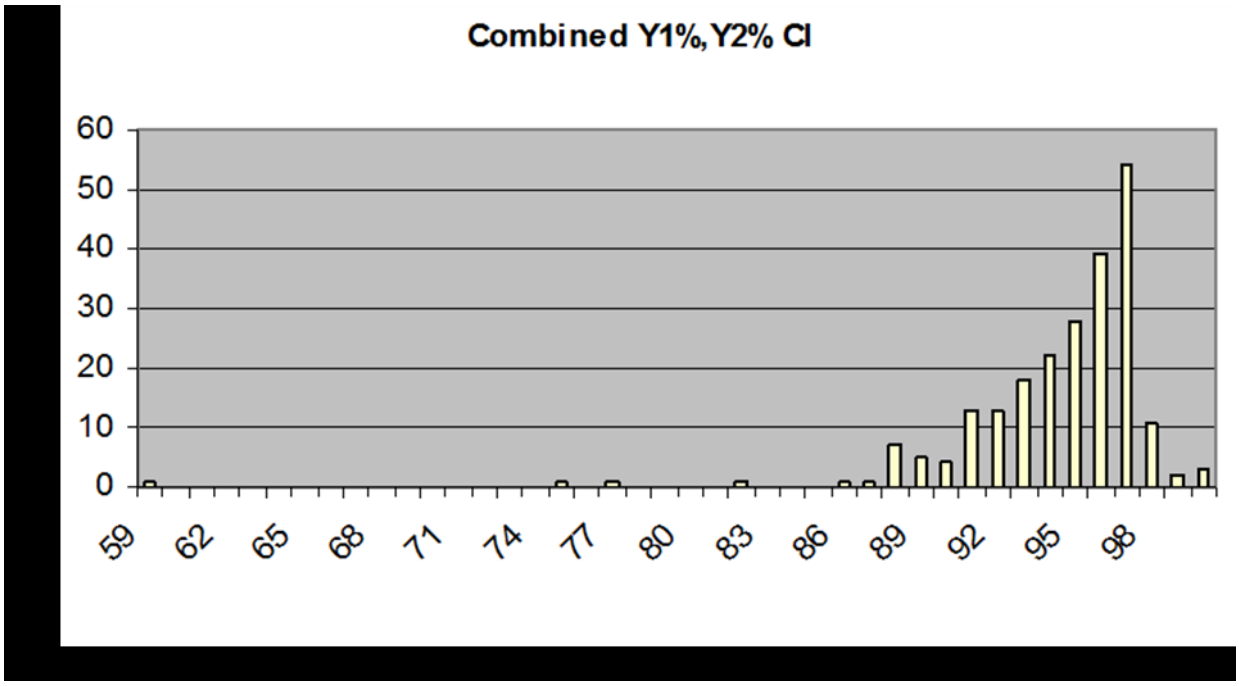

Figure 11: Histogram of the Conscientiousness Index Scores shown as percentages of the maximum possible score, for Year 1 and Year 2 students combined.

	Year 1	Year 2	Year 1 and 2 Combined %
Number of students (n)	110	110	220
Mean (%)	93.53	94.77	94.15
Mode (%)	97	97	97
Standard Deviation (%)	3.68	4.31	4.04
95% Confidence Interval	7.21	8.45	7.92
Kurtosis			27.29
Coefficient of Skewedness			-0.704

Table 10: Table of the Conscientiousness Index Scores.

Shown as percentages of the maximum possible score, for Year 1, Year 2, and Year 1 and 2 combined.

Year	Group	"Happy"	"Concerns"	"Don't Know"	Mean % CI
1	Top	**65**	**0**	**25**	**97.9**
	Middle	**54**	**8**	**28**	**94.3**
	Bottom	**44**	**22**	**24**	**85.1**
2	Top	**69**	**3**	**18**	**98.5**
	Middle	**61**	**1**	**28**	**95.7**
	Bottom	**30**	**45**	**15**	**86.1**

Table 11: Table showing the summed responses for faculty estimates of professionalism for Year 1, and Year 2 students.

Year	Group	Mean	Standard Deviation
1	Top	**6.7**	**1.2**
	Middle	**4.7**	**1.8**
	Bottom	**2.2**	**4.7**
2	Top	**6.6**	**2.0**
	Middle	**6.0**	**1.3**
	Bottom	**-1.5**	**5.6**

Table 12: Table showing the 'Professionalism Index' calculated from responses for faculty estimates of professionalism for Year 1, and Year 2 students.

The results (separately and combined) therefore represent a negatively skewed (towards the right) leptokurtic (values clustered close to the mean) distribution with a major peak between 98 and 99. The similar shape of Year 1 and 2 graphs strongly suggests that there is a similar property underlying these distributions in each of the two years, despite the differences in the components that make up the CI in each year. (For instance, immune status and criminal records declarations are made only at entry into the programme at Year 1).

6.4.2 Validity 1: Relationship between Conscientiousness Index scores and staff views on the trait of professionalism

The responses in each category were added, and the summary outcomes are shown in Table 11. Comparing the "happy" and "concerns" categories by chi-squared analysis shows statistically significant differences ($P > 0.001$ for both year groups). Combining both year groups, 67 out of 79 expressions of concern were found in the lowest-scoring group, 9 in the middle group, and 3 in the top group. The values obtained for the PI (described in the Method section) for each year and each group of students are shown in Table 12.

6.4.3 Correlation coefficients between scores on the Conscientiousness Index and the Professionalism Index

These results relate to averages from the three samples from each cohort. As explained earlier, to further explore the continuous nature of the relationship, a PI was calculated for each student as the “happy” score minus the “concerns” score, and the Spearman Rank correlation coefficient calculated between CI and PI. There was a statistically significantly positive correlation between the CI and PI in each year (0.603 for Year 1 and 0.587 for Year 2).

It was notable that students in the bottom 10 in each year who received the great majority of the expression of concerns also received a significant number of “happy” scores. This indicates that staff views of these students are mixed, with significant numbers of staff having no concerns about them. Staff views of students in the middle and top groups were much more consistent and favourable. Individual results cannot be presented, in view of risks to anonymity, but are consistent with this description. A number of outlier results relate to students with declared disabilities. Such students were frequently perceived as showing good professionalism, even though their CI scores were low.

6.4.4 Validity 2: Relationship between Conscientiousness Index scores and completion of a critical incident report

For Year 2, there were 10 completed critical incident reports. Of these, 7 (70%) occurred in those 10 students who were ranked in the bottom ten in terms (a group representing 9% of the cohort). All instances in which more than one critical incident report was associated with a particular student lay in this group, but further details cannot be given because of the risk of breach of identifiability.

For Year 1, there were 4 critical incident reports recorded, three of which occurred in the 10 students with the lowest CI scores. One of these students received 2 critical incident reports, the only instance of multiple forms being completed in the year group. No further comment can be made on multiple instances because of the risk of breaching anonymity.

6.4.5 Reliability

When the academic year split into two, and performance on the first half year was compared to the performance on the second half year, the Spearman Rank Correlation Coefficient was found to be 0.60 for Year 2 and 0.59 for Year 1, which indicates reasonable consistency for an educational effect.

6.5 Discussion

It can be seen that there is a relationship between the scalar, objective trait of *conscientiousness* as described in this study and the property, or construct, of *professionalism* as perceived by the nine staff members' independent judgments of the students. A total of 67 out of 79 expressions of concern were found in the lowest scoring CI group, 9 in the middle CI group and 3 in the top CI group over the two years together. Calculation of the correlation coefficient between the CI and a summary professionalism index statistic derived by subtracting the "concerns" score from the "happy" score shows that there is a positive correlation, but it is less striking than the comparison with the CI and "happy" scores. This is because the relationship is plainly non linear – as can be seen in Table 2, there is relatively little difference between the top 10 and the middle 10 in terms of expressions of concern, but a marked difference between the bottom group and the middle and top groups There also appears to be a relationship, not yet quantifiable, with the frequency of critical incident reports, particularly in the Year 2 student group. Together, these findings suggest there is concurrent validity for the measure of professionalism through the measures of conscientiousness this study has used. This is in line with previous studies[1-5] that suggest that similar behaviours to those measured here have either concurrent or predictive validity for future practice. It is argued that the CI has the benefit of including a wider range of measures of conscientiousness than have been used before, thus increasing its validity and probable reliability.

Students about whom some staff expressed concerns also frequently received indications that other staff were happy with their approaches to professionalism. This shows that student behaviour is not consistently negative across encounters in different contexts and

with different staff members, in line with previous findings. It also may show the subjective nature of staff members' evaluations—what concerns one staff member may not concern another.

The presence of outlier scores of students with declared disabilities (an outlier meaning that the student has a low CI score but very good staff ratings for professionalism) suggests that the impact of their disabilities was not adequately captured by the CI calculation, and students with declared disabilities should be treated separately in further studies or summative use of the CI.

It is noteworthy that the "don't knows" are highest in the middle group in each year. Staff in this survey frequently spontaneously expressed the view that they would know the really good students and the really poor students in each year, but might not know the "average" students so well. This was confirmed by the observations. Staff also expressed the view that they knew the Year 2 students better than the Year 1 students, through longer exposure, and this view was also confirmed. These results also suggest that the data has validity.

That measurement by the CI is reliable is suggested by two observations. First, the distribution of scores for Year 1 and Year 2 is virtually identical in terms of shape and statistical parameters, suggesting that the same trait is being observed in each year. Second, there is a good correlation between each half of the year, even though the number of CI points differs between the half-years, as the number of occasions for recording

changes. Higher total scores were recorded by Year 1 students because there were more scorable events in Year 1 (associated with the process of induction, such as bringing photographs to induction sessions).

6.5.1 Limitations

This study has a number of limitations. It was conducted in a single institution and there may be context specific factors which affect the conclusions that can be drawn from it. The numbers are too small to explore statistically some of the relationships, notably that between the critical incidents report forms in Year 1, and the CI. Finally, while we used a shared understanding of the meaning of *professionalism,* we did not attempt to define that concept or establish criteria which the nine staff members could have used in making their rankings. Nevertheless, it has been demonstrated that in circumstances such as those that obtain at our institution, the CI clearly correlates with individuals' subjective views of what constitutes *professional behaviour.*

6.5.2 Future work

Research on the CI will be continued, and findings will become stronger with longitudinal sampling, which will take a number of years (perhaps even decades) to complete. The staff ratings of professionalism are subject to all the difficulties with which faculty assessment is associated. However, it was felt the results were sufficiently striking that it would be valuable to bring forward the findings and the methodology at this stage.

These observations suggest that the CI measures a scalar objective trait, and that this corresponds to, or is a valid surrogate for, the construct of *professionalism*, however defined, when the CI is used by experienced staff familiar with the students. Since the individual decisions making up the CI are objective, the measure also has a high degree of inter-rater reliability. It also has the advantage that the data are relatively easy and uncontroversial for administrative staff to collect, as only a record of Yes/No decisions is required (such as whether an individual student has or has not attended compulsory sessions or submitted work) and does not depend on value judgements. The CI could therefore be used as a surrogate measure of professionalism in summative contexts. There are doubts and uncertainties surrounding such a step relating, for instance, to the sensitivity and specificity of the measure, and to setting a cut score for failure. These are resolvable in further studies using the same methodology; most currently employed measures of professionalism do not have this facility. Sensitivity and specificity could be established by comparison with an entire year group over an extended period, when adverse behaviours in clinical settings begin to be reported, while cut scores could be determined by a Contrasting Groups exercise. In the meantime it could be suggested that this measure might be used initially to detect students whose behaviours require investigation and challenge at an early stage in their professional development, and that it may subsequently also be of value in measuring the effect of strategies to help students improve. The findings of this study suggest that in encouraging desirable professional behaviour, targeting students' conscientiousness might be a good place to start.

Part 2: Peer estimates of professionalism and their correlation with Conscientiousness Index scores

6.6 Background

6.6.1 Context

Measures of professionalism in undergraduate medical students are generally subjective in nature, and based on limited observations of behaviours in observed settings. An objective scalar measure of conscientiousness over many occasions, has previously been described, which showed that the Conscientiousness Index correlates with independent faculty estimates of students' professionalism. Part two of this study tests the hypothesis that these measures of conscientiousness relate to independent peer estimates of professionalism, and explore the relationship of conscientiousness with gender and educational background.

6.7 Introduction

As described in part one of this chapter, good measures of cognitive knowledge and practical skills in undergraduate medical education are now available [231]. However, measures of professionalism are much more difficult to implement, stemming in part from the difficulty of defining professionalism, but also from measuring behavioural traits consistently [232]. Key problems are incompleteness of surveillance (students are not observed continuously), lack of accurate reporting (both peers and staff may not honestly report their perceptions for various reasons), dissembling by students (who may role play

expected behaviours) and the differences between teaching environments and real practice environments. However, it has been shown that postgraduate disciplinary action through State Medical Boards is associated with concerns expressed during undergraduate training [233-235], and that low professionalism ratings during Residency programmes are a risk factor for future disciplinary proceedings [236]. This suggests that efforts to measure professionalism systematically and objectively during undergraduate programmes are well worth making, either in order to target remediation, or if that fails, to prevent progression.

Current approaches [112, 232] to measuring professionalism frequently rely on complex and subjective decisions based on occasional surveillance of students. These are of low reliability, lack demonstrated validity, and are time consuming (hence expensive) to implement. They also have low defensibility, in that they are open to challenge by students. However, previous studies have suggested that a lack of diligence, reliability or conscientiousness may be a common feature in predicting later professionalism problems in clinical practice [147, 237]. As a result, a scalar measure of conscientiousness (the Conscientiousness Index; CI) has been developed, which draws on many objective observations of circumstances in which students may choose to be conscientious or not. These include attendance at compulsory sessions, or compliance with essential administrative tasks such as submission of immunisation information. It has previously been found [225] that there was a statistically significant relationship between lack of conscientiousness represented by low CI scores, and staff perceptions of lack of 'professionalism' (as defined by the General Medical Council through "Good Medical Practice" [238]), thereby providing evidence of the concurrent validity of this measure. Reliability was explored through splitting the academic year into two, and comparing individual student performance in each half year, and statistically significant positive and

moderate correlation were observed, indicating reasonable reliability. Students with low CI scores were also those who were most likely to raise staff concerns about their attitude to professionalism, and to be the subject of Critical Incident Reports. In order to establish the sensitivity and specificity of such a diagnostic tool, one requires a 'Gold Standard' comparison. Since this is not available, it is necessary to look for a further measure of validity, which is different from the staff evaluations previously used.

Another approach to measuring professionalism is to seek peer views [113]. Peer assessment, the process of having members of a group judge the extent to which their fellow group members have exhibited specific traits, behaviours, or achievements [115], is a valuable source of information about medical students' professionalism [239]. Its advantages and disadvantages are well documented in the literature[115, 118, 122, 232, 240-245], and are described subsequently. There are three methods of peer assessment in common use: *Peer Nomination*, where students are asked to nominate a limited number of classmates, who best fit various characteristics; *Peer Rating*, where students are asked to rate peers on a series of characteristics according to a fixed scale [115, 241, 243], and *Peer Ranking*, where students rank every peer from best to worst on one or more factors [241, 243]. Peer nomination was chosen for this study. This method, described by Arnold *et al*, [117], is the most frequently studied method of peer assessment [115]. Peer nomination effectively discriminates extremes and it does so with a high degree of validity and reliability[115]. Its drawback is that it is often regarded as being 'virtually useless' as a feedback device because it provides no explanatory information on the substantial proportion of students who will receive no nominations [115]. Although peers have a wider range of observational opportunities than faculty, and are able to witness less 'censored' behaviour on the part of students, evidence suggests that these observations are most

honest and reliable when the stakes are low [232]. As peer assessments begin to have a potential impact on career progression, so collusion between students begins to increase and honest reporting decreases [147]. Nonetheless, peer assessment in formative settings provides a different and independent means of verifying the validity of the CI as a measure of professionalism as compared to faculty assessment. Here the outcomes of such a study are reported. The opportunity was taken to look at the stability of the CI from year to year as data were available, and of the increased sample sizes to compare males and females, and school leaver with graduate and other entrants.

6.8 Methods

This study was conducted at a U.K. medical school during the Phase 1 Programme (years 1 and 2). CI data were gathered over the course of academic years 2006-7 and 2007-8, and is ongoing. Peer evaluation data collection took place over a 4 week period at the end of academic year 2007-08. Ethical permission for this study was confirmed by the Ethics Committee of the School.

6.8.1 Recruitment

All Phase 1 medicine students (93 first year and 112 second year students) were invited to participate in a peer assessment project. An invitation was sent via email and posted on notice boards and via the Virtual Learning Environment (VLE). Invitations outlined the project, and included a copy of the consent form and participant information sheet for study prior to the teaching session in which data would be collected. All students were invited to

email or arrange a face-to-face discussion with the researchers to address any queries or concerns regarding the process.

6.8.2 Consent

All students were issued with an information sheet outlining the purpose of the project, data protection and how any findings would be disseminated. An individual consent form was signed by all participating students. Students were informed that non-participation would not incur any negative consequences.

6.8.3 Method of peer assessment

This study uses a peer nomination approach where students were asked to nominate peers who most or least fit their perceptions of professionalism.

6.8.4 Ballots

Ballot papers were issued to all students at the start of a compulsory lecture. Separate sheets were given to each year group, listing all students within their year. Each student was asked to identify the peer who most displayed their idea of professionalism with an 'M' and least displayed their desired qualities with an 'L'. Students were provided with copies of the General Medical Council's 'Duties of a Doctor'[11] at the time of the ballot as a point of reference for professional (and hence unprofessional) behaviour. Students who did not wish to nominate their peers could still receive 'most' and 'least' nominations from their cohort.

The nominations were compiled and anonymised by two researchers who do not teach on the undergraduate programme. Participants were made aware that faculty would not have access to this data.

6.8.5 Conscientiousness Index

Conscientiousness Index data were collected as described in part one of this Chapter. Recording of conscientiousness points occurred from October to April of the academic year 2007-08. Students were advised verbally and via the VLE at the start of the academic year that this would be done. Students were aware that the CI had no formative or summative role in this iteration. Total conscientiousness points were calculated for each student in both years. Scores were then made anonymous by administrative staff (using a coding system which matched that used for coding peer assessment nominations). Year 2 students had a CI score for the previous academic year 2006-07, which was also utilised in this study.

6.8.6 Categorisation

For the purposes of statistical analysis students were categorised by sex and by the qualification they held on entry on to the medicine programme. Qualifications were classified as: school-leavers (SL): students entering straight after school with A-levels or Scottish Highers, postgraduates (PG): students entering after completion of a previous degree (Bachelors or higher degrees), others (OT): students entering after completion of an access course or the University Foundation programme or holding an International Baccalaureate.

6.8.7 Statistical Analysis

As the peer ranking data were not normally distributed, as determined by inspection and performance of the Kolmogorov-Smirnov of each year group data, nonparametric tests were used whenever this data were analysed statistically. The CI was normally distributed, as determined by Kolmogorov-Smirnov statistical tests of normality in SPSS.

Spearman's Rho test was used to assess correlation between students' CI points scored and the number of nominations gained for the 'most professional' and 'least professional' peer nominations (SPSS v15.0 for Windows). Students who did not receive any votes were not included in the statistical analysis. The data for any students who had a registered disability were also excluded from the data set.

To determine if the proportion of peer nominations between the sexes was equal to the proportion of males and females in the cohort, a test of single proportion was used (Clopper-Pearson test, StatsDirect). To determine if the proportion of CI points between the sexes was equal to the proportion of males and females in the cohort a Student's independent *t* test used (SPSS v15.0 for Windows).

To assess whether there were any differences between the 3 qualification groups for 'most professional' and 'least professional' peer nominations a Kruskal-Wallis test was used

(SPSS v15.0 for Windows), followed by a Mann-Whitney U (SPSS v15.0 for Windows) where appropriate, to determine where any differences lay.

To determine if the proportion of peer nominations (for both most and least professional) between the 3 different groups of qualifications was equal to the proportion of students with those qualifications in the cohort a test of single proportion (Clopper-Pearson test) was used (StatsDirect).

One-way ANOVA (with associated Tukey post-hoc tests; SPSS v15.0 for Windows) was used to assess whether there was any difference between qualification groups for the number of CI points scored. A Bonferroni correction was used to account for multiple comparisons, with α set at $P = 0.05$.

Pearson's Rank test (SPSS v15.0 for Windows) was used to assess any correlation between CI points over 2 consecutive years for the same cohort of students.

In all cases $P<0.05$ was considered statistically significant.

6.9 Results

6.9.1 Participation

118 students in total participated (58%). Of those who submitted nominations 116 cast both ‘most’ and ‘least’ votes. Two participants only cast either a ‘most’ vote or a ‘least’ vote. The percentage of total students in both cohorts that received votes for ‘most’ professional was 57% (117/205). Likewise, the percentage of total students receiving votes for ‘least’ professional was also 57% (117/205). For a further breakdown of the demographics please see Table 13.

6.9.2 Peer nominations and CI points

There was no correlation between students receiving the most nominations for ‘most professional’ and their CI points for years 1 ($p = 0.223$, rho = 0.128) and 2 ($p = 0.197$, rho = 0.123) and both year groups combined (rho = 0.123). However there was a statistically significant correlation between students receiving the most nominations for ‘least professional’ and their CI points for years 1 ($p = 0.002$, rho = -0.314) and 2 ($p = 0.006$, rho = -0.257) and combined ($p = 0.03$, rho = -0.151). These represent a small to moderate negative effect size.

6.9.3 Effect of sex on peer nominations and CI points

There was no significant difference between males and females in the proportion of nominations received for both ‘most’ and ‘least professional’ (Table 13) in years 1 and 2. In addition there was no significant difference in the mean number of CI points between males and females in years 1 and 2 and in both years combined.

	Number of students in year	Proportion of class (%)	Number of 'most' professional nominations received	Proportion of 'most' professional nominations received (%)	Number of 'least' professional nominations received	Proportion of 'least' professional nominations received (%)
Year 1						
Male	55	59	31	72	30	71
Female	38	41	12	28	12	29
Total	**93**	**100**	**43**	**100**	**42**	**100**
OT	16	17.2	6	14	7	16.7
PG	34	36.6	16	37	15	35.7
SL	43	46.2	21	49	20	47.6
Total	**93**	**100**	**43**	**100**	**42**	**100**
Year 2						
Male	52	46	36	49	35	47
Female	60	54	38	51	40	53
Total	**112**	**100**	**74**	**100**	**75**	**100**
OT	20	18	23	31.1	21	28
PG	20	18	27	36.5	2	3
SL	72	64	24	32.4	52	69
Total	**112**	**100**	**74**	**100**	**75**	**100**
Both years combined						
Male	107	52	67	57	65	56
Female	98	48	50	43	52	44
Total	**205**	**100**	**117**	**100**	**117**	**100**
OT	36	18	29	25	28	24
PG	54	26	43	37	17	14
SL	115	56	45	38	72	63
Total	**205**	**100**	**117**	**100**	**117**	**100**

Table 13: Demographic data for year 1, 2 and both years combined.

OT, 'other' qualification group; PG, postgraduate qualification group; SL, school leaver qualification group.

6.9.4 Qualifications and peer nominations

No significant differences were found in the mean number of nominations for 'least professional' between the 3 different qualification groups for year 1 and year 2 students and in both year groups combined. However, there were significant differences between the 3 qualification groups for the number of 'most professional' nominations for year 2 (P = 0.04) and both year groups combined (p = 0.03), but not year 1. Further analysis showed the differences are between the postgraduate students and the school leaver students (year 2, P = 0.01; both year groups combined, P = 0.008). This difference arose in year 2 because the school leavers received 32% of the nominations for 'most professional' but make up 72% of the year group, and postgraduates received 36% of the nominations for 'most professional' but only make up 20% of the year group. In both year groups combined school leavers received 38% of the nominations for 'most professional' but make up 56% of the year groups combined, and postgraduates received 37% of the nominations for 'most professional' but only make up 26% of the year groups combined (Table 13).

6.9.5 Qualifications and CI points

There were no differences between qualification groups for the mean number of CI points awarded in the year 1 group and with both year groups combined. However in year 2 there was a significant difference between the qualification groups and the mean number of CI points awarded (p= 0.004). A Tukey post hoc analysis showed these differences to be between the 'other' qualification group (169.2±0.18; mean ± S.E.M.) and the school leaver qualification group (173.6±0.19; p = 0.006). The effect size was calculated using Cohen's d (d=0.86), which implies a strong effect, but the standard error is significant (0.27).

6.9.6 CI points over 2 consecutive academic years

Since a cohort of students can be followed forward from year 1 to year 2, it is possible to explore how stable the CI is for individual students. A statistically significant correlation was observed between year 1 and year 2 scores ($p = 0.01$, with $R = 0.5$;).

In addition, the changes in CI points (following normalisation of the data) for individual students between Y1 and Y2 was calculated (mean±S.E.M. = 2.13±0.17) and plotted as a histogram (Figure 12). The data shows a normal distribution, with a slight positive skew, showing an improvement in CI points over the two academic years.

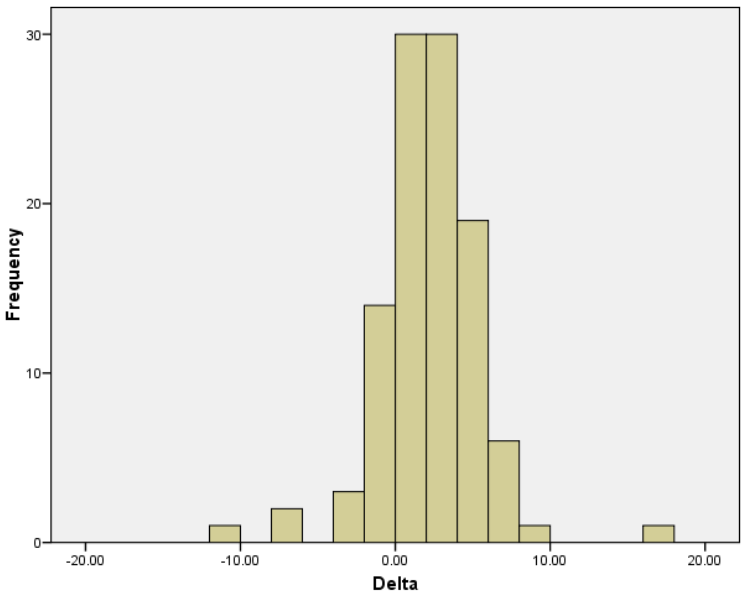

Figure 12:The change (delta) in normalised CI points (percentage of maximum points available) from 2006/07 academic year to 2007/08 academic for the same cohort.

6.10 Discussion

I have previously established [225] that there is a relationship between conscientiousness as measured by a 'Conscientiousness Index' (CI) of undergraduate medical student performance over the course of the entire academic year, and staff views of the construct of professionalism (as defined by the General Medical Council). Here I have demonstrated a significant relationship between peer estimation of professionalism and the CI, providing further support for the concept of the CI as a potential tool for exploring professionalism. This relationship in both studies is strongest between low scores for conscientiousness and concerns about professionalism, rather than between high scores for conscientiousness and positive views about conscientiousness – in other words, a negative correlation between the CI and perceived professionalism. This suggests that while staff and peers find it difficult to be precise about the presence of professionalism, they nonetheless have some common understanding of the absence of professionalism. This could be paraphrased as "I know when I don't see it." It could simply be that bad behaviour is more memorable.

The correlation coefficients are low to moderate in Cohen's terms [246] as effect sizes go in education. However, this is comparable with the effect sizes observed by Papadakis *et al* [233, 234]. This may relate to the finding of Papadakis *et al* that expressions of concern during an undergraduate programme are sensitive but not specific as a diagnostic test for future difficulty. If there is a degree of change in individuals, then this may explain a component of the variance, and stable low scores may be significant for future predictions. Colliver and colleagues [247] criticise Papadakis for 'exaggerating the importance of the relationship between unprofessional behaviour in medical school and subsequent board action', and deem the link between the two to be 'too weak to have any prognostic

value'[247]. However, it is not unusual to introduce medical treatments on the basis of such figures, recognising that the 'Number Needed to Treat' may be high. More significantly, disciplinary action by state boards is only one measure of poor performance as a clinician. Arguably, it is both rare and extreme. In order to calculate the true predictive value, we would need to identify the number of professionalism lapses that did not result in board action such as drug dose errors, missed diagnoses and failure to follow up cases. Some but not all of these data are in principal obtainable. In any case, we would not wish to suggest that conscientiousness is the sole trait or construct involved in professionalism, but rather that these figures suggest that it is a significant contributor. Conscientiousness is therefore likely to be of significance both in measuring and in modifying professionalism in educational settings.

I also explored the relationship between CI in year 1 and year 2, and found that there was a statistically significant positive correlation between scores for individual students over two years, suggesting that the CI has at least some stability. While the observed correlation of 0.5 is high as an educational effect [246], it is perhaps lower than expected as a measurement of a stable trait. If there is a degree of change in individuals, then this may explain a component of the variance, and stable low scores may be significant for future predictions. In general there was a small increase in percentage performance from year 1 to year 2, and if this is confirmed in other settings and years, may provide a measure of changes in professionalism.

Finally, I explored the relationship between CI and demographics. Perhaps unexpectedly, there was no significant difference between male and female students or between graduate

students and school leavers. This was unexpected, as generally, females are regarded as being more diligent than males. This assumption has been made within the university as well as being an assumption frequently expressed on a societal level. Significant differences were observed between school leavers and students defined as 'other', but the 'other' group is both small and heterogeneous and I would not wish to over interpret this finding at the moment.

While the relationship between the CI for undergraduate students and professionalism in future practice can only be identified over a long time period, the existence of the CI data allows such a long term relationship to be explored. However, in the light of the evidence that undergraduate performance is related to future performance in terms of professionalism, the CI provides an objective scalar measure of undergraduate performance which seems to be relevant. In the meantime, there is the possibility that the CI can be used to monitor performance and to target interventions. Its major advantages over other approaches are the objective nature of scoring, the continuous nature of the observations, and the low cost of gathering the data. Of course, a key question is whether or not CI could be used as a determinant of progression, and this cannot yet be answered. However, one possible summative use may relate to acceptance of remediation. There is some evidence that students can be both incompetent and unaware [248], and it may be that failure to remediate after interventions for low CI could be used in subsequent progression decisions.

Participation rates and comments on participation (which will be considered in a separate qualitative study) confirm that there is a reluctance on the part of students to rate each

other, in line with the literature on this subject[241] . In general they did not report professionalism as being important to them at this stage.

6.10.1 Limitations

A limitation of this study was the lack of participation from some (particularly year 1) students. There were concerns that students might vote by popularity, awarding ‘most professional’ votes to friends, and students that they disliked ‘least professional’ votes. More generally, this study was conducted in a single institution and would require repetition in other settings to explore how generalisable its use might be.

Chapter 7: Peer and self-assessment of Consientiousness

Part 3: Peer estimates of Consientiousness Index Scores within tutor groups

7.1 Background

Parts 3 and 4 of this study aimed to ascertain whether there was a correlation between students' self and peer perceptions of professionalism and Conscientiousness Index scores. Furthermore, I sought to determine whether any correlation between the CI and peer assessment was enhanced by greater familiarity with those peers assessed by students. Literature [3, 4] suggests that the reliability of peer assessment is improved by the number of assessors and the number of observations. This study assesses peers within a small tutor group in order to investigate this notion.

7.2 Method

7.2.1 Ethics

Ethical approval was granted by the School of Medicine and Health Ethics Sub-Committee.

7.2.2 Recruitment

Year 1 & 2 undergraduate medical students in the 2008-09 cohort were invited to participate. Recruitment procedures followed those outlined in Chapter 2.

7.2.3 Data collection

A peer nomination exercise was delivered on the VLE. Tutor groups, each consisting of 10 students, were invited via a seperate email to participate in this study. The email contained a link to an individualised exercise for each group. As in part 2 (Chapter 6), students were required to nominate the peer they perceived to be 'most' and 'least' professional from their cohort. The two major differences between this part of the study and part 2 (Chapter 6) were:

1. Students could only nominate peers from their tutor group.
2. Students were given the option to provide a free-text justification as to why they nominated each peer.

The exercise was created using the Blackboard™ survey tool. Students selected their chosen 'most' and 'least' professional peers from a drop down menu. Free-text comment boxes were provided underneath each of the drop down menus.

7.2.4 Quantitative data analysis

The author downloaded student responses from the VLE using Excel (for Windows 2003). The nominations for each student in the cohort were collated. The CI scores for each student were provided by administrative staff and were inputted with nominations into Excel. Once nominations and CI scores were coupled for each student all data were anonymised, students were coded using their anonymous Z –number which is assigned upon course registration. A Kolmogorov-Smirnov test was then performed to check the data for a normal distribution.

7.2.5 Qualitative data analysis

Free-text comments (vote justifications) were collated using Word 2007 for Microsoft Windows. These free-text comments were collated, for year 1 and 2 separately.

Comments were categorised according to 'most justifications', 'least justifications' and 'explanations of non-participation'. The comments were then coded by the author (GF). These qualitaitve data were coded inductively,using grounded theory, as per the method descirbed in Chapter 2.

7.3 Results

7.3.1 Quantitative results

Data were analysed using a Pearson's correlation as data were deemed to be parametric (SPSS version 15.0 for Windows). Results are summarised in table 14.

Year	'Most' (rho)	'Most' (p)	'Least' (rho)	'Least' (p)
1	**0.514**	**0.000***	**-0.292**	**0.003***
2	**0.248**	**0.018***	**-0.275**	**0.008***

Table 14: Results of the Pearson correlation between 'most' and 'least' votes for each year group with Conscientiousness Index scores.

* denotes a significant result (p>0.05).

7.3.2 Qualitative results

The themes for 'most justifications', 'least justifications' and 'explanations of non-participation' will be reported in turn. Year 1 and 2 results are reported separately for each category.

Year 1- Justification of 'most' professional nomination

Seven reasons (themes) emerged from the data describing why year 1 students nominated a particular peer as most professional. These were attendance, punctuality, contributions, professional views, behaviour in the clinical context, preparedness and dress code.

Each reason will now be described in turn, with exemplary quotes for each.

Attendance

Students were likely to be nominated as most professional if their peers had noted them to have a good attendance record.

“I think she had been to every session...”

Punctuality

Nominators placed a strong emphasis on the punctuality of their peers.

“Punctual to sessions.”

*“...*and is always on time....”

Contributions

Students who actively contributed, or whose contributions were deemed to be relevant were voted as most professional. The context from which these contributions arose appeared to be that of the small group tutorial, where students debate ethical, professional and political issues influencing health.

“Has valid opinions and even if they are against the majority view, voices them.”

"...when she does she always has professional views and often says things that I hadn't even thought of. However, she never jumps in or is overpowering when she contributes, allowing others to have their say too. She is friendly and easy to get on with, and she doesn't participate in the 'joking' that goes on within the group and which I feel hinders our discussion."

"He is friendly and participates actively in discussions."

"One of the only people to make a decent contribution to the group."

"Ethically aware in discussions."

Professional demeanour

Nominating students perceived those they voted to have a professional demeanour.

"She always speaks in a professional manner."

"Behaves in a professional manner on academic visits [to hospital]..."

Clinical conduct

Students who behaved appropriately in a clinical environment were praised by their peers.

"Most professional due to age and healthcare experience."

"...has worked in a professional hospital environment before and thus has a lot of experience dealing with patients."

"Anyone would happily see X in a consultation."

"... conducts herself well in clinical settings ..."

Preparedness

Students were likely to be given votes for being most professional if they were well prepared for teaching sessions.

"She is always prepared for the tutorial sessions."

"He's always prepared for work..."

Dress code

Those students who were appropriately turned out impressed their peers.

"...dresses well for hospital visits and just generally appears to be professional..."

"...is always well turned out."

Year 1-justifications for 'least' votes

Justifications for allocation of a 'least' vote fell into 5 categories; joviality, discriminatory attitudes, attendance, ill-preparedness and inappropriate dress.

Joviality

Voting students were not accepting of peers who exhibited unacceptable levels of joviality.

"...laugh too much - he plays the joker- and although I know that it has been this light-heartedness that has allowed the group to relax and become more comfortable with each other, I think it should stop when the session officially begins."

Discriminatory or unprofessional attitudes

Least votes were allocated to those who expressed attitudes which were causes of concern.

"Often expresses un-professional beliefs, and acts inappropriately in professional situations e.g. on a visit to a GP surgery when placed with a HCA he commented that it was useless going with her as he knew more than her already and it would be a waste of time."

"Inappropriate comments that show a discriminatory attitude."

"...and say things that sound a bit judgemental."

Attendance

Poor attendance resulted in a least professional nomination.

"...rarely turns up to sessions...."

"She always misses sessions."

Ill-preparedness

A lack of preparation for tutorials and examinations was criticised by peers.

“...she often admits that she hasn't done any study (for last set of exams) and appeared comfortable with this. I would be mortified to think like that. Sorry.”

“He almost never reads any material for the tutorial before a session (often he doesn't have it with him either)...”

Inappropriate dress

Inappropriate clothing was not tolerated by nominating peers.

“She often does not dress appropriately when we are sent on visits.”

“...clothing not always suitable.”

Year 1: justifications for non-participation.

Some students were not willing to participate in the peer nomination exercise. Instead, they chose to provide an explanation why they declined to participate.

There were too few explanations to identify any meaningful themes. Some example comments have been reported instead.

“This kind of survey is completely unjustifiable, and smacks of the GMC/HPC's penchant for getting people to "rat on their mates". All members of the group are professional, and there should be no distinguishing between them. If there is a matter of unprofessionalism, there are channels through which it can be dealt with, and it is not the place of fellow students, or anyone else, to randomly assign a grade of "professionalism.”

“I had to select somebody. This is an awful question, and encourages a negative attitude towards those selected. Everybody in the group appears professional. It's like asking a Mother who is her least favourite child!”

Year 2: Justifications for ‘most’ votes

Year 2 justifications for awarding ‘most’ votes fell into the following themes: ‘personable’, ‘dress-code’, ‘contributions’ and ‘intelligence’.

Personable

Nominating students awarded ‘most’ votes to those peers they deemed to be personable.

“Highly personable.”

“X is generally outgoing and pleasant to work with.”

“Approachable, friendly.”

Contributions

Students who made significant contributions to tutorials were perceived as being professional.

“X always makes valuable contributions to group discussions, he is keen to put his point across and does so in a constructive way. In addition to this he is keen to listen to the opinions of others, he is open minded and sensitive towards others.”

“X speaks with substance and confidence (his views always add value to the group discussion), and respects other people's opinions...”

Dress code

An appropriate appearance contributed towards students receiving a ‘most’ professional vote.

“...and always looks presentable.”

“X is smartly dressed for all hospital visits.”

Intelligence

Interestingly, nominated students were deemed by peers as being high academic acheivers.

“ He is also very clever and understands the content of the course very well.”

“Always scientifically knowledgeable.”

“X is very brainy.”

Year 2: justifications for 'least' professional votes

Three major themes emerged from the justifications provided. These were 'punctuality', 'egocentric views' and 'inappropriate dress'.

Punctuality

'Least' votes were awarded to those students who were persistently late.

"X is friendly enough, but has missed sessions, is often late..."

"X tends to be unpunctual and does thing last minute."

Egocentric views

Students were critical of peers who could not accept the views of others.

"Discourteous in group sessions (when he attends). Quick to put down others opinions, and rarely listens to others viewpoints without butting in!"

"Is not punctual nor good at seeing things and understanding other points of view."

"Makes inane and irrelevant comments about issues unrelated to the areas of discussion. Argues her personal opinions and is un-open to those of others. Or if she listens she

conveys herself in a manner to ensure all present are aware she does not agree nor accept another's opinion."

Inappropriate dress

Nominators were swayed by the clothing of their peers.

"X doesn't dress for the situation."

"[Least professional] only in terms of professional dress and communication skills".

Year 2: justifications for non-participation

As with year 1, there were not enough comments regarding non-participation to produce an in-depth analysis. Exemplary comments are provided below:

"I found this difficult, as I feel everyone in my tutor group is professional."

"Due to everyone having different personal qualities, this is very hard to gauge."

"No one in the class is "least professional". Again this is question is poor, unfair and reflects a badly though and unprofessional form of peer review. I agree with the GMC that peer review is very important during our medical education, but feel strongly that the methods employed by Durham are unsuitable. The quotation used in the email from the GMC document has been taken out of context as reveres to peer review of students as teachers

and mentors. If you have any evidence base that this is appropriate I would be happy to read it."

7. 4 Discussion

The Pearson's correlations have shown that greater familiarity with the assessees enhances the correlation between professionalism as measured by peers (votes) and professionalism as measured by faculty (CI scores).

Peer assessment within tutor groups correlated significantly at both the most and least professional end of the spectrum, respectively, for year 1 ($p<0.001$, $p=0.003$) and year 2 ($p=0.018$, $p=0.008$).

This study therefore further validates the use of conscientiousness, as the student outliers as identified by the Conscientiousness Index were also identified as either cause for concern or exemplary by their peers, as measured by votes received.

It could be speculated that greater familiarity with those being assessed improved the reliability of the peer assessment exercise. Basing the assessment on tutor groups enabled the number of observations potentially considered to be increased as the assessors and assessees attended multiple compulsory teaching sessions together each week, as well as attending clinical and community placements together. This meant that students were observing a wider range of behaviours displayed by their peers by observing those they

assessed in a variety of contexts. This array of behaviour is often not witnessed by faculty when conducting assessments of professionalism, as students are often able to 'switch on' professional behaviour in the presence of those they believe to be their superiors or responsible for assessment.

These findings highlight the importance of familiarity upon the validity of an assessment. The implications of this are that institutions wishing to conduct peer assessment should ensure that the assessor has adequate knowledge of those they are assessing, within the correct context, in order to ensure that the assessment outcome is valid. This is especially important if the assessment is for summative purposes. That being said, this study was formative. Replication of this study design where the peer assessment was for summative purposes may have a very different outcome. This assumption is based upon literature which states that assessments where participants deem the outcome to be low-stakes are more valid [97, 121, 249].

The justifications for peer nominations provided by both cohorts of students offer not only an insight into why students nominated their peers, but also into their definitions of professionalism.

In Chapter 8 students' views on professionalism are described, these were opinions collected during focus groups. As described previously (Chapter 2), focus groups can be easily dominated by more vocal participants. The resulting opinions may therefore be socialised, and not participants' true feelings. The free-text justifications coded in this study

were collected anonymously and in isolation. Data may therefore be more representative of the participants' true opinions.

The justifications provided by year 1 students were more broad ranging than those provided by year 2. It is plausible that the year 1 students have more diverse opinions regarding what constitutes professional behaviour. It is worth noting that the participation rate for year 2 was significantly lower than year 1 and may therefore explain why the opinions of year 2 participants were less diverse.

Punctuality and attendance featured heavily as reasons for both positive and negative nominations. Students weighted tutorial contributions and conduct within the tutorial discussions highly. Votes for "least professional" were often attributed to egocentric students, who were dominant or unable to accept views which differed from their own.

As will become evident in Chapter 8, students placed an emphasis on clothing. In Chapter 8 descriptions of a professional doctor are of someone who dressed the part. The justifications provided in this chapter also emphasise dressing appropriately as an important attribute for a medical student. It is unfortunate that the reasons behind this emphasis were not provided.

From a speculatory perspective, students are socialised at medical school. This is with respect to behaviours and attitudes. Students are requested to dress appropriately for the OSCE in the May exam diet. The subliminal message may therefore be that faculty are putting a significant emphasis on clothing. Students may even perceive that faculty regard

appearance more highly than other aspects of professional conduct, such as courtesy. Clothing is an integral part of identity formation[219, 221, 222], as was discussed in Chapter 5 and will be further described in Chapter 8.

Part 4: Further peer estimates of Conscientiousness Index scores within tutor groups

7.5 Conceptual framework

In parts 1-3 of this study, students assessed the professionalism of their peers without standards to base their assessments upon. Part 4 of this study repeats the study design in part 3, peer assessment in tutor groups. This was to ascertain whether the results attained in part 3 were reliable. Scores were also aggregated to identify any relationship between CI scores and student classification as “most” or “least” professional.

7.6 Method

The peer assessment exercise was delivered online using the secure VLE. The exercise was created using Bristol Online Surveys© (BOS)[250]. Students were asked to consider their peer members of their tutor groups; which typically consist of ten students. These groups were chosen as students spend the most amount of timetabled hours with these

specific peers. A survey was created for each tutor group. Students accessed the survey for their group via an individualised web link, accessed through the VLE.

The survey listed the names of the students in the tutor group, excluding the name of the student completing the survey. Students were requested to select, from a drop down menu, the name of the student they deemed to be 'most' professional and the name of the student they deemed to be 'least' professional within their tutor group. There were also two free-text boxes to enable the participant to justify their "most" and their "least" nomination. Consent was taken to be completion of the survey.

Data were downloaded and subsequently collated using Excel (for Windows 2007). Nominations for "most" and "least" professional from each group were combined into one spreadsheet, thus the cohort was analysed as a whole. The number of least votes for each student was deducted from their most votes to produce an 'aggregate score'. These data were all correlated with the CI scores.

Statistical analysis was conducted using SPSS version 16.0 for Windows. A Kolmogorov-Smirnov test was performed to test the data for normality. As data were parametric, they were analysed using a Pearson correlation.

The variables analysed were: 'number of most votes' or 'number of least votes' and 'aggregate score' with 'Conscientiouness Index score'. The results are presented in table .

After establishing whether a correlation existed, the data were further analysed. Students were ranked according to aggregated scores. Students were categorised according to their aggregate score. Categories were ‘positive’, ‘neutral’ or ‘negative’. For each category the mean CI score for the students was calculated (Excel 2003 for Windows). The standard deviation and standard error of the mean were also calculated (Excel). Results are reported in tables 15 and 16. The procedure was also repeated for the precious cohort (2008). The aggregate scores are shown for the 2008 and 2009 cohorts in table 16 and figure 13.

7.7 Results

Participation was 66.6% (68 out of 102 students) for year 1, and 44.1% (41 out of 93 students) for year 2. In total, 55.9% (109 out of 195) of those invited participated.

The Conscientiousness Index scores correlated with ‘most’ and ‘least’ professional votes for both cohorts. Aggregated scores were also correlated with CI scores. All correlations were statistically significant ($p<0.05$).

Year	Most versus CI	Least versus CI	Aggregate versus CI
1	R=0.464 **p<0.001***	R=-0.324 **p= 0.001***	R=0.497 **p<0.001***
2	R=0.298 **p= 0.004***	R=-0.228 **p= 0.030***	R=0.325 **p= 0.002***

Table 15: Results of Pearson correlation.

Statistical significance was taken to be p<0.05. * denotes a statistically significant result.

Cohort	Year group	Number of students	Aggregate score	Mean CI score	SD	SEM
2008	**1**	7	Negative	187.0	5.7	2.1
		55	Neutral	191.2	4.5	0.6
		23	Positive	193.1	4.8	1.0
2008	**2**	10	Negative	169.5	4.6	1.4
		63	Neutral	172.7	5.7	0.7
		31	Positive	173.7	4.9	0.9
2009	**1**	20	Negative	173.6	7.5	1.7
		42	Neutral	176.3	5.5	0.8
		31	Positive	179.6	5.5	1.0
2009	**2**	23	Negative	161.4	3.3	0.7
		40	Neutral	162.5	3.7	0.6
		28	Positive	163.8	1.9	0.4

Table 16: Mean CI scores for categories of aggregate scores.

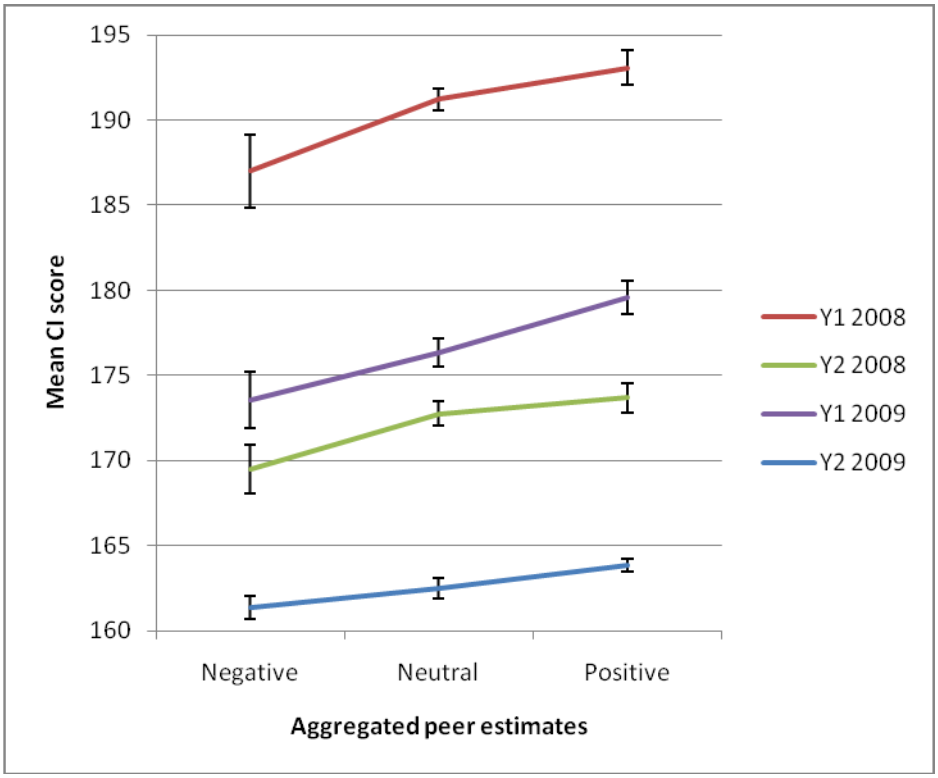

Figure 13: Aggregated peer estimates and the corresponding mean CI scores for each cohort

7.7.1 Discussion

As parts 3 and 5 discuss peer estimates of conscientiousness in detail, I will only briefly discuss the results of part 4 of this study here. It is worth noting that the peer assessment results achieved in part 4 of this study mirror those from part 3. This demonstrated that the accuracy of peer estimates of professionalism were enhanced by greater familiarity between peers. This is likely to be due to the increased number of opportunities for peer observation when utilising assessors in the same tutor groups, a finding which mirrors the literature[116, 120, 121]. The correlation achieved in parts 3 and 4 of this study demonstrate that the relationship between professionalism, as judges by peers, and faulty

measures of the construct of professionalism are consistent. Results may allude to conscientiousness and professionalism being stable traits.

Additionally the mean CI score for each of the aggregated peer estimate scores demonstrated that, on average, those who were scored positively by their peers had attained a higher Conscientiousness Index score than those who were scored neutrally or negatively. This was true for all four sets of data presented (figure 13). The graph (figure 13) shows that the mean CI score attained by students increases in a linear manner as the students' professionalism categorisations improve from negative to positive. This demonstrates a strong ordinal relationship between the construct of professionalism, as judged by peers, and the construct of conscientiousness, for these broad categories ('positive', 'neutral' and 'negative'). The trends in CI scores for students categorised as 'positive', 'neutral' or 'negative' have been further analysed but this analysis is outside the remit of this thesis.

Part 5: Peer and self-assessment of Conscientiousness Index scores using histograms

The rationale behind part 5 of this study was that I wanted to determine whether, when presented with Conscientiousness Index data, students were able to accurately assess their own placement and the placement of their peers within the cohort distribution. Although parts 1-4 explore professionalism, part 5 explores conscientiousness as there is no formal professionalism assessment within Durham University by which all students could be compared. Althought the constructs of professionalism and conscientiousness are different, the process of making judgements on performance is similar.

7.8 Method

Due to teaching, and thus the possibility of students collecting additional CI points, this part of the study was conducted after the term had ended. CI scores for the academic year 2008-2009 were collated for each year group. For each year group a frequency distribution of CI scores was produced, which were subsequently represented as histograms (Excel 2007 for Windows).

This second exercise, also created using BOS©[250] was accessed using the VLE, as described previously. Students were shown the histogram for their cohort. Despite the

histogram being representative of the cohort, students were required to only consider their tutor group. Each student was then asked to select which bin on the histogram they believed that each member of their tutor group, including themselves, had attained a score. Students were also given an option to provide a free-text justification of each placement they made.

For each student the mean Conscientiousness Index score, as determined by their tutor group, was calculated. The self-assessment score for each student was also isolated.

In order to determine if there were any relationships between the scores allocated by the department (attained CI scores) from peers or from the self-assessment, each bin on the histogram was converted to a CI score. As the bins represented a range of scores, the median value was given for each bin. For example, a student who self-assessed that they lay within the bin range 155-160 would be given the score 157.5.

The peer and self-assessed scores were then correlated with the CI scores students attained across the year, as determined by the department. In order to do this all scores for the entire cohort were collated and analysed. Data were analysed using a Pearson's correlation as data were deemed to be parametric by inspection and Kolmogorov-Smirnov test. Data were analysed using SPSS version 16.0 for windows.

7.8.1 Qualitative data analysis

Free-text justifications from the assessment exercise were collated in Microsoft Word. Data were analysed using grounded theory, as described in Chapter 2. Each cohort's data were analysed separately. However, themes were consistent therefore results are reported for both cohorts combined.

7.9 Results

7.9.1 Quantitative results

Participation was 79% (81 out of 102 students) for year 1, and 27% (25 out of 93 students) for year 2. In total, 54% (105 out of 195) of those invited participated.

Year 1 students were unable to self-assess CI scores ($R= 0.145$, $p=0.241$), but were able to accurately place their peers within the CI distribution ($R= 0.373$, $p<0.001$).

Likewise, year 2 students were unable to self-assess CI scores ($R= -0.055$, $p=0.814$), but were able to accurately place their peers within the CI distribution ($R= 0.452$, $p<0.001$).

7.9.2 Qualititative results

Ten major themes emerged from the free-text justifications. These were:

Attendance, participation in focus groups, submissions, professional conduct, completion of strand evaluations, position as student representative, assisting with widening participation activities, contributions to teaching sessions, punctuality and MedSoc contributions.

Each theme will be described in turn with exemplary quotes.

Attendance

Histogram positioning was dependent upon students' attendance.

"I have not missed many lectures and have never been late to a [tutor] group…"

"X was absent couple of times for the [tutor] sessions."

"…he always attends lectures and seminars."

"Unprofessional, misses lectures, falls asleep in lectures he attends and is often late."

Attendance at compulsory sessions

Students specifically attributed CI scores to attendance at compulsory classes.

"Turns up to most compulsory sessions…"

"Attends all lectures - both compulsory and non-compulsory…"

Participation in focus groups

Placement in the CI distribution was also reliant upon a student's particiaption on voluntary foucs groups which staff conducted as part of research projects.

"I have done a few interviews and focus groups."

"I have not participated in a large amount of research [focus groups] as a subject and I think that this will have prevented me from gaining extra points."

Submissions

Perceived punctuality with regard to assignment submission was also a consideration.

"… X handed in all assignments punctually."

"I know I have submitted assignments etc on time…"

Professional conduct

Professional conduct was considered when allocating students to the appropriate histogram bin.

"Very professional."

"X is a fantastic all rounder, he is thoughtful, supportive and does all and more at Uni."

"I feel that my peers are more professional and organised than I am."

"Hard-working and diligent so likely gained a lot of points."

Position as student representative

Peers who volunteered to be student representative of student-staff consultative committees were credited when scores were considered.

“X was very active this year, being a student representative.”

“ X was a student representative.”

Completion of strand evaluations

Completion of end of strand reviews was a factor in assigning CI scores to both peers and self.

“I also missed two or three evaluations.”

“I completed most of the unit evaluations.”

“X always fills feedback forms in.”

Assistance with widening participation activities

Students are invited to volunteer to assist with outreach activities, geared towards widening participation in Higher Education. Contributions to this programme were acknowledged.

“I have participated in voluntary activities.”

“I've participated in a few voluntary activities that are usually rewarded with conscientiousness points.”

“X helps with all the school visits...”

Contributions to teaching sessions

Students considered the contributions of themselves and their peers during teaching sessions.

“X puts a lot of effort into all the areas of the programme and is also very dedicated.”

“…contributes well in [tutor] sessions...”

“I feel sometimes that X can be a little bit apathetic about the course and doesn't really put in as much effort as the others.”

“X always volunteers in class and participates fully.”

Punctuality

Punctuality was considered when allocating CI scores.

“I attended each group session on time.”

“Always on time, not late.”

MedSoc contributions

Students who undertake voluntary positions of the Medical Society (MedSoc) committee were rated highly.

“X is the MedSoc xx.”

“I help with MedSoc…”

7.10 Discussion

The quantitative results of this study show that year 1 and 2 students were unable to self-assess their conscientiousness ($p=0.241$; $p=0.814$). However, the same students were able to accurately assess their peers’ conscientiousness ($p<0.001$; $p<0.001$).

The ability to peer assess has previously been demonstrated (Chapter 6 and part 3 of this Chapter). The interesting result is the inability of students to self-assess; this may demonstrate a lack of self-awareness. It would appear from the correlation coefficients ($R= 0.145$, $R=-0.055$) that the second year students were less self-aware than the first year students.

This lack of self-awareness replicates the findings of Kruger and Dunning[251]. This famous paper reports how their participants in a self-assessment exercise tended to hold overly favourable views of their abilities in many social and intellectual domains. Kruger and Dunning suggest that “this overestimation occurs, in part, because people who are unskilled in these domains suffer a dual burden: Not only do these people reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the metacognitive ability to realise it[251].”

Kruger and Dunning[251] argued that incompetence causes poor performance but also it causes the inability to recognise that one's performance is poor. They consistently found across the four studies that participants, scored for various tasks, in the bottom quartile not only overestimated themselves, but thought they were above-average. The results presented in this chapter have demonstrated ‘unskilled and unaware’ in a real-world, higher-stakes and practice related setting.

Boud (1999) [252] defines a self-assessment as “the involvement of students in identifying standards and/ or criteria to apply to their work and making judgements about the extent to which they have met these criteria and standards.” The results presented here may be indicative of students’ inability to identify the levels of conscientiousness they have exhibited over the course of the academic year.

Although the participating students did not receive the results of the assessment exercise, making students aware of the results of this self and peer assessment of conscientiousness may serve to illustrate that you cannot rely on self-assessment and that peer assessments are valid insights which offer a useful perspective for the individual concerned to consider. Educators should highlight to students that external assessments are legitimate means of receiving feedback.

The inaccurate self-assessments could be down to many factors, for example, inaccurate recall, modesty or arrogance . Students may have selective memories, only recalling situations where they perceive themselves to have done well. My findings regarding

conscientiousness are unique, however, they mirror the literature in that students are consistently poor at self-assessing [253-255].

Self-assessments are judgements, based upon intrinsic and extrinsic factors. These factors might include one's self-perception or eccentricities, a history of events, or influences from others, as well as many other unconscious factors. This could be likened to reflexivity in qualitative research; when coding qualitative data using the grounded theory paradigm the researcher acknowledges that they have presuppositions about the data and/ or the research topic of which they may or may not be aware. I propose that this is the same concept for an individual conducting a self-assessment exercise and that as educators it is our duty to enable students to consider what their intrinsic biases are so that we can steer students towards making assessments which are less influenced by presuppositions.

Despite the consensus within the literature that ability to self-assess is generally poor[252-255], self-assessment is not a wasted exercise. Articles advocate self-assessment as a key step in professional development[253-256]. Completion of both self and peer assessment exercises promotes reflection and thus self-awareness. This alone is beneficial to students as they continue on a path of professional and personal development. Such opinions and self-judgements will drive the students' learning. Furthermore, exercises such as these acclimatise students to both giving and receiving feedback. The emotional response to completing an assessment on one's self or peer, coupled with receiving feedback, is necessary for learning. Boud[252] argues that feelings and emotion are involved in learning, and that any reflection would be inhibited if "these dimensions are not accepted as integral to it."

As I stated earlier, in this study I have attempted to demonstrate the Kruger and Dunning[251] “unskilled and unaware” phenomenon in a real-world setting. Demonstrating this phenomenon within the medical school context is important as throughout their careers medical students will be required to self-assess. This will be both on a formal and informal or self-motivated basis. Self-assessment may form one arm of a multi-source feedback exercise or could be an individual deciding to do further reading after identifying a gap in their knowledge. Both examples require the subject to make an informed decision about their own ability to perform on some level. As illustrated by Eva and Regehr [255], the term self-assessment has been used to describe a variety of processes which include self-directed assessment seeking, reflection, self-regulation and self-monitoring. Other studies[257] described self-audit and self-rating. While Eva and Regehr[255] may argue that using the term in such an all encompassing manner is troublesome, I would argue that whatever the context of the assessment, the principle remains the same; by going through this process the student is prompted to make a judgement on their own abilities. This judgement may not be accurate but the path followed is one of self-monitoring, reflection and regulation. These are skills are imperative for life-long learning.

Personal and unguided reflections do not provide the learner with enough information to adequately guide their performance improvement[255], and thus I would argue, their learning. Perhaps by coupling self-assessment with peer assessment, as in multi-source feedback or in a simple form as in this study, we might be able to provide students with a context in which to consider their own performance. By considering the performance of others, students may begin to identify their own misperceptions about their own abilities. A

systematic review[257] of the accuracy of physicians' abilities to self-assess concluded that physicians have a limited ability to self-assess and that "the processes currently used to undertake professional development and evaluate competence may need to focus more on external assessment."

With regard to the qualitative data presented, justifications of histogram placements were focused. Students appeared to have a clear perception of what constituted conscientiousness, whether this was their own perception or their version of the faculty measured CI is unclear. The considerations made by students mirrored the faculty CI point awarding system in many ways, as described in Part One. The award system is known to students but, as previously described, is deliberately vague and is not reinforced. This consistency with faculty measures is surprising since, anecdotally, students claim not to understand the Conscientiousness Index.

Factors considered by students can broadly be categorised as: diligence (attendance, punctuality, submissions etc), additional activities (focus groups, student representatives, MedSoc) and demeanour (professional conduct).

These categories shed light on students' perceptions of professionalism, specifically from the pre-clinical (years 1 & 2) perspective. Caution must be aired, as students may be regurgitating their perceptions of faculty definitions and expectations. Nonetheless, the justifications offer an insight into the students' understanding of this complex and subjective topic.

Justifications were made for both peers and self. Students' justifications showed a significant level of reflection, specifically self-reflection and self-awareness.

As we will see, the distinction between the rationale for awarding professionalism scores in this study and the opinions on professionalism expressed in Chapter 8, is the non-clinical focus. In Chapter 8 students clearly describe professionalism as being relevant in the clinical context, with little emphasis on the university setting. In this study the justifications were entirely based upon academic events, such as assignment submission and lecture attendance. Students received no guidance on what to consider when allocating histogram bins, so the option to factor in behaviour on clinical visits was available. I imagine that the absence of the clinical context will largely be down to awareness of the Conscientiousness Index criteria. However, these academically focused observations are useful as they highlight to faculty the aspects of professionalism students perceive to be relevant within the university context. This is fair to say as, anecdotally, the majority of students are unaware of the technicalities of the CI.

It is worth noting the participation rates in this study were significantly lower for year 2 students. From my speculative perspective, this could be a lack of willingness for year 2 students to peer assess and their perception that participation risk relationships with peers. By year 2, students have often forged life-long friendships and may perceive that participating in a peer-assessment exercise could jeopardise the relationships which are important to them. Of course, low participation could be attributed to any number of factors such as timing, concerns over anonymity, or use of data, to name but a few.

Chapter 8: Student's views on professionalism

8.1 Background

Part One of this study has been presented at international conferences[258, 259]. Part Two of this study has been published in the journal Medical Education[260] and has also been recorded as a podcast for the journal .

8.1.1 Contributions to this chapter

I was responsible for the study design for both parts of this chapter, analysed data and wrote the manuscripts submitted for publication.

Dr Marina Sawdon contributed to the data analysis for both parts of this study. She also contributed to the manuscript for Part Two of this study, published in Medical Education[94]. Ms Jayne Garner contributed to the data analysis for Part Two of this study. She also contributed to the manuscript for Part Two of this study, published in Medical Education. Dr Laura Clipsham was involved in the data collection and analysis of Part One of this study. Professor John McLachlan commented on the manuscripts submitted for publication, which consequently form part of this chapter.

8.2 Introduction

Professionalism is subjective and neither easily defined, nor quantified[225]. Definitions of professionalism are extensive and disparate, with most authors listing numerous and diverse attributes, for example Van De Camp *et al*.[261] list 90 attributes compiled from a

systematic literature review. Likewise there are many guidelines on professionalism, with a number of professional guidelines being issued to students when they commence their studies, for example the GMC's "Good Medical Practice"[262] and "The Duties of a Doctor"[263]. How much understanding students have of these guidelines, and their perceived relevance, is little known. It has been reported that medical students know the professional behaviour that is expected of them and can imitate this in an exam[264]. The importance of informal and hidden curricula in shaping student perceptions of professional behaviour and values have also been highlighted as valuable learning experiences[265], as has the impact of role modelling[266] .More recently the impact of social networking has been associated with unprofessional behaviour by medical students[267] and raised issues relating to appropriate behaviour on websites such as Facebook[268]. Students can develop rhetorical strategies to cope with unprofessional situations such as dissociating themselves from the identity of the qualified professional and associating themselves with the identity of the student, and switching between the two; the so-called *identity mobility* strategy[269, 270].

Students' views on professionalism are well documented in the literature [117, 271-274], but few describe the opinions of Phase 1 students[275]. Phase 1 medical students are those in the first two years of their five year medical degree. An understanding of how Phase 1 medical undergraduate students perceive professionalism may help us teach professionalism more effectively at this early, and formative, stage.

Peer assessment, often used for assessing medical professionalism, has been defined as the process of cohort members judging the extent to which their fellow students exhibit

specific actions. These can be traits, behaviours or achievements[115, 117]. The unique peer perspective has been noted by many[112, 117, 121]; this is because peers witness routine behaviour, rather than the modified or cautious behaviour often displayed when classmates are being directly observed.

The reliability of peer assessment is dictated by the number of observations, observers and items being observed [121]. Other factors which have a strong influence on the reliability of peer assessment are the relationships between peers, whether "high stakes" exist, and the nature of the assessment (formative or summative)[121], as well as the assessment design[112, 115]. Peer assessment is established as an assessment tool within clinical practice, and is now utilised extensively in undergraduate assessment of professionalism.

Part One of this investigation was conducted in succession to a study[276] validating a informative measure of professionalism, the Conscientiousness Index[225], through use of peer assessment. Subsequently, students participated in this qualitative study which sought to determine student's perceptions of professionalism, and how they thought peer assessments ought to be conducted to optimise reliability, and participation.

Part Two of this investigation sought to determine students' perceptions of professionalism, and how they thought peer assessments ought to be conducted to optimise reliability, and participation. Our previous research[92, 225, 258, 259] was therefore the basis for our conceptual framework. I sought to provide the perspective of UK, pre-clinical medical students, thus filling a gap in the literature. Part Two of this study reports on one main

cluster emerging from these focus groups, students' perceptions of professionalism and the context in which it is relevant to them.

8.2.1 Reflexivity

As this chapter is qualitative I will now describe how I view the topic of professionalism. My interest in medical professionalism stems from researching the Conscientiousness Index[276]. However, a real passion for the subject was as a consequence of my attendance at an event hosted by the University of Liverpool's Centre for Excellence in Teaching and Learning (CETL). Here I listened to a presentation by Professor David Stern, who described how peer assessment was utilised in the United States as a measure of professionalism. The result was this study correlating peer assessment and Conscientiousness Index scores published in Medical Education[276]. After conducting the quantitative study[276], I sought the opinions of students on both professionalism and peer assessment, the subject of this chapter.

The background to this study is important, as I was highly influenced by the work of David Stern. I had definite preconceived ideas about students' acceptance of peer assessment, and how they might wish such an assessment to be conducted. Although this might have biased the data collection, my focus group questions were as a result of these presuppositions. Using a ground theory approach to data analysis meant that coding was inductive. I allowed the data to generate the theory, and tried to the best of my ability to let both sides of the argument emerge, where existent. My training as an anatomist, to my knowledge, had no bearing on this study.

Other influences with regard to data analysis were the opinions of colleagues involved. Dr Laura Clipsham is medically qualified and thus has an interest in any comments which were clinically oriented. Ms Jayne Garner has a background in local government, and thus an interest in policy, this is reflected in her interest with GMC guidelines and the students perceptions of them. Jayne was also employed by the University of Liverpool CETL and her research foci included professionalism. Dr Marina Sawdon, at the time, had no background in medical professionalism research. She was therefore the least influenced by presuppositions.

Part One: a pilot study of students' views on professionalism and peer assessment

8.3 Background

Previous studies suggest that students' views on professionalism placed importance on knowledge and skills, relationships, lifestyle, culture and family[277]. Studies often report high proportions of students being opposed to participating in peer assessment[116, 117, 121], but it is still recommended as a formative assessment tool for assessing professionalism in medical undergraduates.

Part One of this study describes how undergraduate medical students perceive professionalism and their views on peer nomination as an assessment tool for measuring professionalism.

8.4 Methodology

Data collection took place over a 1 week period at the end of the academic year 2007-08. Part of this methodology is reported in Chapter 2 and in the publication *Medical Edcuation*[276].

8.4.1 Ethical approval

Ethical approval was granted by the University's School for Medicine and Health sub-committee prior to commencement of this research.

8.4.2 Recruitment and consent

All Phase 1 medicine students (93 first year, 112 second year students) were invited to participate. Recruitment and consent procedures were as described in Chapter 2.

8.4.3 Focus groups

Focus groups were conducted to ascertain the students' views of peer assessment as a tool for measuring professionalism, following conduction of a peer nomination process to determine students' views on peers who most and least fit their perceptions of professionalism (this data will be reported in a separate paper). Data were collected through focus groups as this method allows group discussion facilitated by the authors in order to explore multiple and opposing opinions.

Groups were conducted until thematic saturation occurred. Twelve students participated; therefore 3 focus groups were conducted, each with 4 students. One focus group was first year students only (1 female, 3 males); the other 2 were second year students (1 female, 3 males in one group and 2 females, 2 males in the other group).

The author facilitated semi-structured focus groups. Dr Laura Clipsham also facilitated focus groups. The focus group discussions were recorded and transcribed verbatim. Transcripts did not identify participants in any way. All students were given the opportunity to approve transcripts before publication.

8.4.4 Data analysis

Data from focus groups were analysed using a grounded theory approach. Open coding was carried out by authors independently. Axial coding of the themes was conducted by authors, with negotiation. This procedure is described in detail within Chapter 2.

8.5 Results

During the coding process two main themes clearly emerged: students' perceptions of how professionalism relates to them and students' views of the design of a peer assessment tool. Relating to student views on professionalism, the key subthemes which emerged were perceptions of attributes of professionalism (both positive and negative), relevance of professionalism to students and teaching & learning of professionalism. In relation to peer assessment, 5 subthemes emerged: the use of an online environment for voting, eliminating anonymity, receiving prior warning, being able to justify choices and participation promoting reflection. These themes are depicted in Figures 14 and 15.

1. Perceptions of professionalism

Perceived positive attributes

When reflecting on peers' professionalism participating students considered enthusiasm for medicine, participation in extracurricular activities, willingness to help others and strong study habits to be the key attributes of professional behaviour.

"...enthusiasm, outside lectures for the subject, doing extra reading, outside work not just the stuff that was set, promptness..., attendance and involvement with extra curricular activities."

"…it's based purely on...study habits because... that's all... that's really important at the moment... someone I aspired to work in the same way as..."

Perceived negative attributes

Attributes considered to be unprofessional were less diverse. Participants' main concern was for peers displaying unprofessional behaviours towards guest & clinical lecturers.

"…[unprofessional behaviour is] *rudeness and lack of courtesy to guest lecturers"*

Relevance of professionalism to students

Participating students perceived themselves principally as “students” rather than young professionals, expressing the belief that professionalism did not need to be displayed in lectures nor whilst studying. They expressed a clear desire to be permitted to retain the freedom traditionally associated with being students, and talked of “getting away with it”.

“...despite what the GMC say about us being doctors now,… I like to think of myself as a student so in that sense we're still able to get away with things that obviously we won't later on… in my mindset I'm taking advantage of that while I still can...”

“...we are students so we shouldn't have to be that professional in lectures but when you've got external visitors in… then you should be on… better behaviour than maybe you'd get away with with everyday lecturers. It's about adjusting to each circumstance I think.”

Furthermore, participating students often disregarded the applicability of professional standards outside of the clinical setting.

“...it’s really important to be able to make that switch yourself, like I act like this outside the hospital, I act like this in the hospital, and as long as they manage to act professional[ly] in the hospital, I personally don’t have that much of a problem with it as long as when they’re in a clinical setting they know how to act...”

Teaching and learning

Participating students felt that receiving teaching on professionalism is beneficial in order to raise awareness of what was expected and provide a contextual basis to their learning. They strongly expressed the desire for guidance on what is expected of them at each stage. They felt this would prepare them more for the transition from medical student to junior doctor.

"...there's got to be... some kind of transition from being a medical student to a professional person in a clinical area and I think we should have a bit more direction as to what is expected..."

Participating students' preference was for signposting of professional behaviours as opposed to professionalism being taught didactically, many deeming this inappropriate as they considered professionalism to be innate.

"I don't think you can learn a lot of this professionalism stuff."

"...it's quite easy to sit down in a lecture hall and say 'this is professionalism' with so many bullet points and this is why it's important, but it's much harder, I don't think that is teaching someone to be professional. Learning about professionalism doesn't necessarily help someone develop their skills of being a professional..."

2. Peer assessment design

Peer voting environment

Participating students strongly expressed a preference for peer assessment to be completed online, rather than in a lecture theatre with the rest of their cohort, to increase participation, decrease peer pressure and allow greater consideration of choice.

"…I think online would work better, you'd feel more comfortable, probably be more inclined to take part."

Anonymity

The use of an online assessment process was associated with anonymity. Although introduction of an online assessment was regarded positively by students, anonymity was strongly opposed by some participating students.

"If you're so worried about someone seeing it you shouldn't be saying it… it's not really right to be putting that opinion forward if you're not willing to stand behind it."

"…not be anonymous and actually tell people in a group or… person to person… it's constructive criticism… and then they could improve…"

Time

All students unanimously agreed that it is imperative to receive prior warning of a peer assessment exercise, as they felt that this enabled thinking time before making final decisions.

Justification

The majority of participating students agreed that peer assessment would be perceived more positively if there was the opportunity to justify their own decision, with a strong emphasis on enabling justification of the choice of student who they perceived to be least professional.

"…it's a bit ...vacant just to put somebody's name... it would be nice to... justify it a bit."

Additionally, it was believed that a lack of justification would increase the likelihood of receiving unfair or inaccurate assessments from peers as it would allow for a greater influence from relationships with peers.

"I think justifications would be important because it would stop people trading off, because if you've got to really justify your answers, you can't really say 'well I picked my best mate'…"

Reflection

Students commented that participating in peer assessment promoted reflection upon their own behaviour.

“I think it makes you reflect on your own actions as well as other peoples.”

“...you start to think about yourself ‘while I'm choosing them for that, how am I?’...”

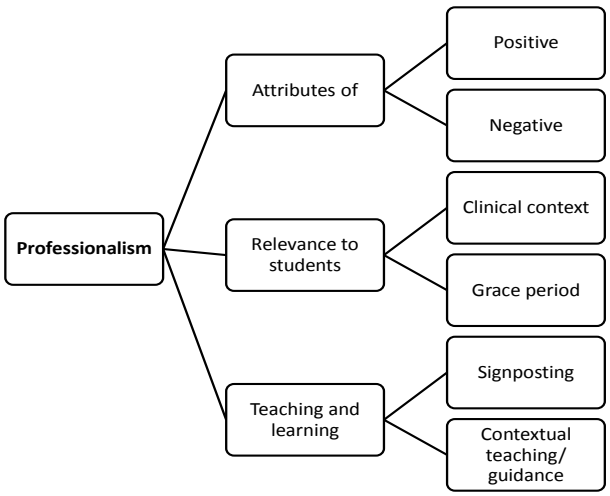

Figure 14: Professionalism themes.

Map showing one of the main themes (professionalism) emerging from focus group discussions (far left box), and the subthemes within this (centre boxes). The boxes on the far right are the descriptors of the subthemes.

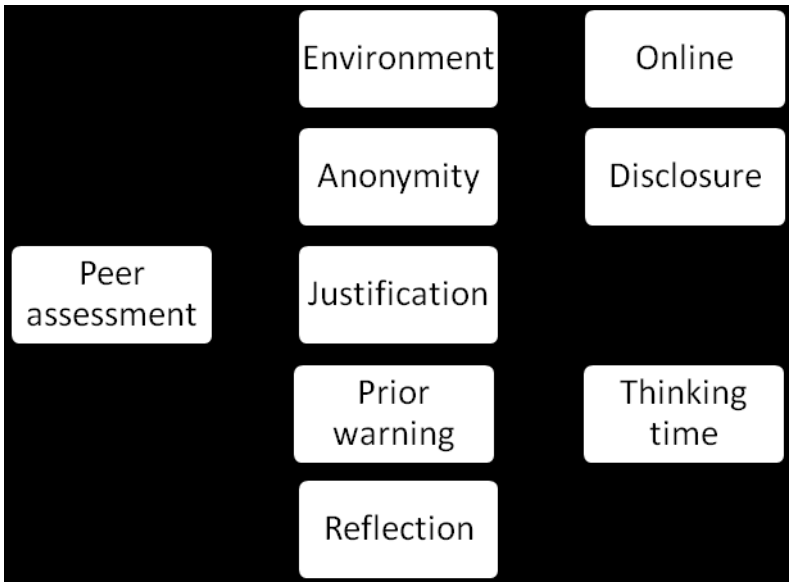

Figure 15: Peer assessment themes.

Map showing one of the main themes (peer assessment) emerging from focus group discussions (far left box), and the subthemes within this (centre boxes). The boxes on the far right are the descriptors of the subthemes.

8.6 Discussion

The two main themes arising from the focus group discussions gave an insight into how students view professionalism as not relevant to them unless they are in a clinical situation, and an overall acceptance of peer assessment as a tool to measure professionalism provided they can justify their choices openly.

Medical students perceive professionalism as applicable in the clinical, rather than the academic context. This extends to showing professional courtesy and behaviour towards clinical and visiting lecturers, whilst feeling they do not need to extend this to academic staff to the same extent. Students have a poor grasp of the professional standards expected of undergraduates, with many suggesting that as students they should be allowed to behave in a more reprehensible fashion than they would as clinicians. This is further supported by their comments regarding changing standards of behaviour towards internal and external (clinical) tutors. Students talk of "acting" professionally in a clinical setting and knowing when to "make the switch". However, what students understand and define as professionalism does not correlate. When discussing their thoughts on professional behaviour a contradiction became apparent; students often listed academically related attributes such as study habits, enthusiasm for the subject and doing extra work beyond that in the curriculum as being professional. This addresses a question posed by Arnold and colleagues[121] whether students use concepts of professionalism related to being a student or those of being a doctor when assessing their peers. It appears they know that professionalism should be shown in an academic situation; however they feel that as students they should be able to "get away with it". This perception of professionalism was not expressed in a study carried out by Wagner and colleagues[277] where students' views on professionalism placed importance on knowledge and skills, relationships, lifestyle, culture and family.

Students in the preclinical years feel that professionalism is nothing to do with behaviour in a lecture nor outside the university, but solely in clinical situations. Professionalism being badged as only relevant in a clinical domain is not surprising, as generally it is taught within this context, by practising clinicians. The students seem focused upon the role of the doctor

and the clinical context when discussing professionalism, but had very little acknowledgement of the patient and none at all of collegiality. This makes comments about behaving professionally to external tutors and not towards existing staff more striking. Professionalism is being viewed as strongly sequestered in particular contexts and not at all in others, going against the guidance from *"Duties of Doctor"[103]* and *"Good Medical Practice."[230]*

These comments are indicative of 'proto-professionalism' as described by Hilton and Slotnick[278] which encompasses the developmental period of professionalism, prior to the students having clinical experience. Without this practical experience the evolution of students' knowledge and skills base is limited. During this "proto-professionalism" phase their behaviour is perhaps determined by moral and reflective judgements rather than by their experiential learning and subsequent practical wisdom[279].

With so many definitions of professionalism and guidelines on expected standards of behaviour in circulation then it is not surprising that students have difficulty defining it and grasping the extent to which it is applicable during all stages of their medical career, including the undergraduate years. As educators, we believe that undergraduate medical students cannot be expected to behave as fully fledged medical professionals, especially during their preclinical education, as they do not yet possess the knowledge and skills required to do so. Students should, however, start to demonstrate some of these behaviours. It would seem that students are making a moral decision not to do so. It appears that there may be a level of resentment from medical students that there is an expectation of them to act in a certain manner that is not expected of most other university

students. This perhaps stems from the vocational nature of the course, of which students are well aware of upon application.

The lack of understanding of the importance of professional behaviour in medical students suggests a shortfall in the teaching and learning of professionalism in the academic context. Students expressed a wish for more direction and guidance on what they see as the "transition" of professional behaviour between the student and the junior doctor. Some students felt that teaching of professionalism was beneficial but several others thought professionalism cannot be learnt and that didactic teaching of professional behaviour does not allow students to develop the required skills and attributes. Throughout our medical curriculum there are currently 3 aspects that are important in developing professionalism in medical students:

1. Staff ethos (role modelling)

2. Mode of curriculum delivery (provision of a clinical tutor to each student to provide clinical mentorship)

3. Curriculum content (from the GMC's "Duties of a Doctor"[103])

The third, curriculum content is thought to be the least important of the three. Perhaps, then, more emphasis should be put on professional behaviour in the curriculum, perhaps in the form of debates as one student suggested. However the difficulties in teaching professionalism are well documented[279-281].

The other main theme was that of peer assessment as a tool to assess professionalism. There was an acceptance of early peer assessment exercises in the pre-clinical years, as students understand that they will be required to do this in their clinical years and thus feel that early exposure to it will give them valuable experience of this process. However, one of the caveats included that it could be accomplished online to increase participation and reduce the pressure of peers being able to observe nominations. In contrast to this, there was an overwhelming desire to abolish the anonymity of the peer assessment as students feel that you should have the courage of your convictions and use the assessment as constructive criticism, allowing the peer to reflect on the result. This is in contrast with the findings of Gukas and colleagues who showed that students prefer anonymity for peer feedback, and that peer assessment made them reflect on their work[282]. Other studies have shown positive and mixed views on anonymity[117]. Students discussed professionalism as a moral virtue, that when you are reflecting upon a peer's professional behaviour you are making a moral judgement.

Being able to justify one's choices in peer nomination of professionalism was important for the majority of participants. There was a feeling that this would reduce the amount of peer pressure to positively nominate one's friends. Indeed, not justifying their choices appeared to decrease the significance of their choices, leading them to feel that they were just picking a name or "ticking a box".

The comments of Arnold and colleagues[116, 117, 121] that the context in which a peer assessment occurs will influence student participation levels, are acknowledged. However,

student comments enabled us to delineate characteristics of a peer nomination assessment of professionalism that students feel would encourage them to participate:

- Substantial prior warning, to allow “thinking time”.
- Online assessment
- Opportunity to justify choices
- Disclosure of nominations and reasons given by peers

Durham students are in agreement with Dannefer and colleagues[107] that it is plausible to introduce peer assessment as a formative measure of professionalism in an undergraduate medical school programme. For this to be successful students must be comfortable with the process and fully understand the judgements they are about to make, and their consequences. The opportunity for peer assessment to act as a tool for initiating student reflection upon their own behaviour is highly beneficial.

8.6.1 Limitations of the study

The limitations of Part One of the study was the small sample size of self selecting participants and the presuppositions of the facilitators. However it was felt that bias was minimised by using facilitators who have no role in the teaching of professionalism. Future research will increase participant size with the aim of confirming students’ perceptions of professionalism in undergraduates.

Part 2: “You’re judged all the time!” Students’ views on professionalism: A multi-centre study

8.7 Background

Part one of this study was submitted to the journal “Medical Education”, but was extended to a larger cross-institution study in response to reviewer feedback.

Editors were supportive of the research and its value and therefore suggested collaboration with another institution. Links were established with the University of Liverpool, making an obvious choice for the extension of Part One of this study. As mentioned previously, this extended study is now published[260].

8.7.1 Context

This study describes how medical students perceive professionalism and the context in which it is relevant to them. An understanding of how Phase 1 students perceive professionalism, will help us teach this subject more effectively. The inclusion of another institution allowed for the findings to be generalised to the UK pre-clinical context as students were not from one culturally homogenous institution.

8.8 Method

8.8.1 Ethics and recruitment

Ethical approval for the study was granted by the Ethics Committees at Durham University and the University of Liverpool. As Lancaster Medical School is a satellite delivering the Liverpool curriculum, the University of Liverpool granted ethical approval. Participation in the focus groups was voluntary. Theoretical sampling was applied to invite students at the appropriate educational level (i.e. Phase 1 medical students following a five year curriculum with supervised clinical contact) to take part in the focus groups. Based upon preliminary research[225, 258, 259, 270], authors from Durham University (GF,MS) collaborated with the University of Liverpool (JG) as a means of triangulating data; this also served to increase the sample size and ensure that any findings were generalisable to the wider UK context. Theoretical sampling [160] allowed authors to gain a deeper understanding of students' views on professionalism in a specific context, that of Phase 1 training. Durham and Lancaster Universities both offer separate Phase 1 training. Other than stage of training (i.e. participants must be in Phase 1), there were no other exclusion criteria for this study.

All students in the Phase 1 cohorts at both institutions were invited to participate in this study. The recruitment and consent protocol outlined in Chapter 2 was employed.

8.8.2 Focus groups

Focus groups were conducted at the two medical sites by the three authors, with transcription shared between the authors and administrative staff at both institutions.

Data were collected through thirteen semi-structured focus groups as this method allows group discussion facilitated by the authors in order to explore multiple and opposing opinions. Groups were conducted until thematic saturation occurred, lasting approximately one hour for each group.

A topic guide was used for the focus groups, derived from previous research by the authors and general literature on professional behaviours and undergraduate medical students [92, 225, 258, 259]. This allowed for a semi-structured approach and minimised facilitators influencing the discussion (see Table 17). The focus group discussions were recorded and transcribed verbatim. Transcripts did not identify participants. Students were given the opportunity to approve transcripts before publication. During the focus groups and data analysis stages authors carried out regular member checking[283, 284] to confirm their interpretation of the data.

Focus Group Topic Guide	
Spine:	**Key points for discussion:**
Professional behaviours	Definitions, attributes
Relevance of professional behaviour	Setting, stage of training, applicability
Teaching, learning and assessment of professionalism	Student opinion: how to teach, learn and assess
General comments	As directed by participants

Table 17: Focus group spines

These are the focus group spines used in order to facilitate discussion on professionalism amongst the Phase 1 students.

8.8.3 Participant profile

This study considers the opinions of students from 2 medical schools in order to ensure that the students were not culturally homogenous. All participants were Phase 1 medical students. The age of participants varied between 18 and 40 years. Phase 1 students at Lancaster were following a Problem Based Curriculum (PBL), whereas Durham students were following a case-led curriculum. Despite Phase 1 being pre-clinical training, students

have a significant amount of clinical contact. Clinical contact was hospital-, community- and classroom-based. Students at both institutions were taught by academic and clinical members of faculty. Following Phase 1 training students from Durham and Lancaster merge with cohorts from Newcastle and Liverpool Universities respectively for their Phase 2 clinical training (years 3 to 5 of medical school).

The students' backgrounds varied. Participants had either entered medical school straight from secondary (high) school, had previously completed a Bachelors or Masters degree, or had practised as an allied health professional. Both institutions had UK and international participants. The participants who attended were representative of the cohorts with respect to age, ethnicity and educational background.

8.8.4 Data analysis

Data were collected and analysed in iterative cycles, and coded using NVivo Version 8 by authors. Data were analysed using a grounded theory approach [283-286] with constant comparison. The thematic coding of the data were inductive and undertaken individually by the authors, who then met to agree content groupings and internal interpretation validity[287].

8.8.5 Reflexivity, micro social theory and social constructionism

With regards to reflexivity[163, 286, 288] the authors in this study had no role in the teaching or assessment of professionalism for any of the participants; this avoided power relationships and any influence over data collected.

A grounded theory approach was utilised as it is both a methodology and a paradigm. Grounded theory acknowledges the possibilities of biases of which an author may not be aware; by using such a method authors allow the theory to develop from the data rather than findings being subject to the authors' biases[158].

As the research focuses upon personal and immediate social interaction of daily activities and personal experience, it can be classified as microsociology[289]. These interactions demonstrate a shared understanding of processes, situations and language from which individuals make sense of their world[290]. As researchers, it is important to acknowledge our own definitions and comprehension of professional behaviour, as this has shaped the questions we are asking students in the focus groups. Therefore, this research takes a social constructionist approach as it does not aim to discover scientific facts, but seeks to explore students' views and experiences of professionalism.

8.8.6 Facilitator and author profile

Authors acted as focus group facilitators for this study. The authors were all experienced in qualitative research and focus group facilitation and had no formal role in the teaching and assessment of professionalism at either institution. With regard to reflexivity, the authors' backgrounds are varied; including physiology (MS), anatomy (GF), and local government (JG). All authors are currently employed within medical education research posts.

8.9 Results

Of the 245 students invited, 72 students (31 males and 41 females), all of which were in Phase 1 of their training, attended focus groups at two institutions. Of the 72 students, 15 were from the University of Lancaster and 57 were from Durham University.

Seven major themes emerged from the data. These were: context, role modelling, scrutiny, identity, ‘switching it on’, leniency, and sacrifice. Each theme (underlined), its subthemes, and mini-themes are described in turn. A diagram of these themes, subthemes and related mini themes, as well as relationships between them, is depicted in Figure 16.

Themes are the most prominent clusters of information to emerge from the data; these are the principal concepts. Within themes, subthemes have been described. These describe conceptually discrete aspects of the theme which have emerged from the data. Mini-themes are smaller concepts which are nested within and explicate the sub-themes, or describe discrete concepts within the overarching sub-theme. The authors acknowledge that ‘mini-themes’ are not a conventional reporting method for qualitative data, however, we have devised this hierarchical coding system to demonstrate the development of ideas within a theme. The hierarchy delineates the discrete, yet relational aspects of concepts within a specific theme.

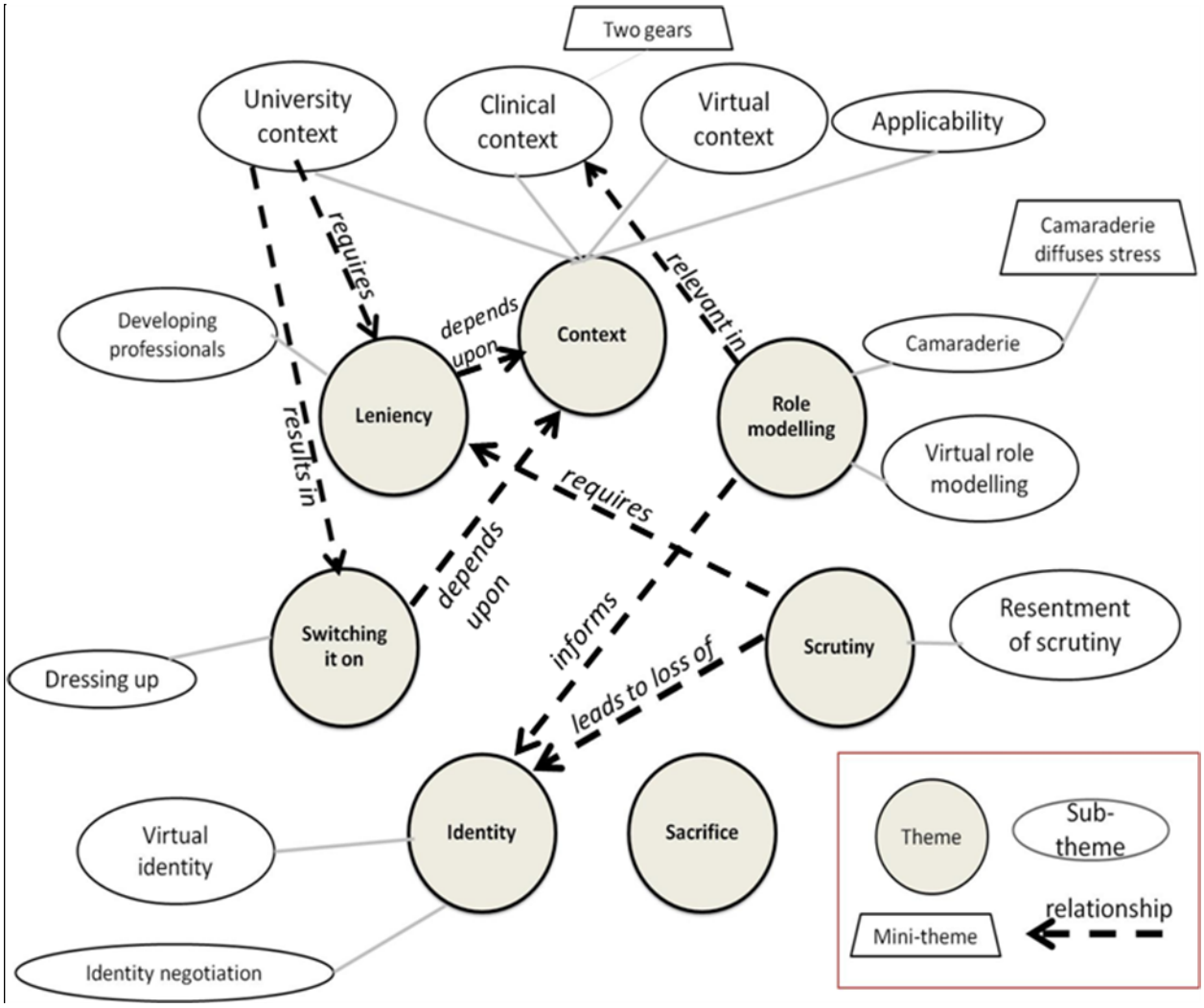

Figure 16: Themes emerging from students' views on professionalism.

Diagram depicting the main themes (shaded circles in the centre), subthemes (ovals linked to the main themes) and mini themes (trapezoids). Dashed line indicates further relationships between themes.

One theme to emerge from the data, which will not be reported in the results, was definitions of professionalism. Students' understandings of professionalism ranged from behaviours, attitudes, values and attributes. This discussion added nothing new to the

literature, nor to the readers understanding of the context of the results outlined below. Many of the definitions provided mirrored those described by Van de Camp et al (2004)[96].

Students also described how they wished a peer assessment of professionalism exercise to be conducted. The view points were consistent with those voiced in the data from part one of this study, and thus are not reported here.

Context

Students described professionalism as being context dependent. Three contexts were described; the university, the virtual, and the clinical.

Clinical context

Students regard professionalism as being predominantly relevant within the clinical context.

"…it's most relevant to me to be professional in a hospital setting…"

'Two gears' of professionalism

The analogy of 'gears' refers to a manual (stick-shift) car. "Two gears" is interpreted as meaning two speeds or levels of behaviour. Within the clinical context of professionalism "two gears" were repeatedly described: professionalism in the presence of patients and professionalism in the absence of patients. Furthermore, these gears were often

anecdotally described as being exhibited by individuals that students perceived to be role models.

"…the medics that we shadow have almost two gear behaviour. There's the banter that they'll have with you [students] and then when the patient has walked in it changes".

"…different degrees, levels of professionalism like around patients you have to be more professional…"

Virtual context

The second context in which professionalism was described was the virtual, with specific reference made to the social networking site 'Facebook'. Students refer to their 'status' which is a brief statement or headline on their personal profile page. This status is often visible to all members of the social networking community.

"...I've noticed that people in medicine seem to constantly have … stuff to do with their course as their Facebook status..."

University context

Students made reference to a third context of professionalism, the university environment.

"...I think lecturers probably would be a bit more forgiving if you were … less professional..."

Contextual applicability of professionalism

There was confusion regarding in which context professionalism was relevant.

"*I don't know what professionalism means in the context of a university student*"

"I don't understand where it [professionalism] matters here [Phase 1]. I don't see where it comes from and where the relevance of it all is"

Role modelling

When discussing ways in which professionalism may best be taught or learned participants advocated observing role models, specifically within the clinical context.

"… observing… doctors…when we're in the hospitals and on visits … that's a good way to see how they behave … and that's like a good role model to try and follow…"

Role models and the virtual context

Lapses in professional behaviour were witnessed by students on social networking sites. Incidents involving role models were described.

"I've got a friend.. she's a doctor now and often I notice her [Facebook] status is things like 'so and so is sick of intravenous drug users' and I'm thinking 'this is awful, you can't put that on Facebook' and I mean she's not like naming names so maybe she thinks it's ok... it's always to do with something that's happened at work."

Camaraderie

There was an association between role models and unprofessional behaviour, excused by students as camaraderie. Anecdotally, students described this camaraderie as 'banter'.

"…it's a necessity to be able to have fun with your peer groups as doctors and, yes, maybe this will involve sarcasm, a bit of joking around and things like that, maybe discussing patients, you know, probably not in the most professional way, between yourselves…"

".. what you're saying in a group of clinicians together, isn't going to be what you're going to say in front of a patient anyway…"

Camaraderie as a stress diffuser

This ‘camaraderie’ was condoned as a legitimate means by which medics can diffuse stress.

“I think you should always be able to relax a bit you know when you’re with your peer group otherwise you just turn into robots...then you would probably fizzle out pretty quickly...”

Scrutiny

Students were aware that their professional behaviour will be scrutinised, both within and out with the clinical environment.

“*Our behaviour outside the surgery will be on scrutiny, you know if you get a skin full and have a fight in a pub as a practising medic, you’re not going to be practising for very long. So in a way, learning to live … within those boundaries is important...*”

Resentment over scrutiny

Not all students were accepting of this scrutiny.

“God you’re just judged all the time!”

“It’s a terrible burden to place on people”.

Identity

Another major theme to emerge was that of identity. Subthemes were identity negotiation and virtual identity negotiation.

Students struggle with the distinction between their identity as a medic and as an individual.

"...there's like a blur between them..."

Identity negotiation

In addition students talk of an identity negotiation, often related to context and environment.

"*I see myself as a medical student so I think that the medical part comes out when I go to hospital and then the student part comes out in like lectures and outside of work so I kind of think I'm allowed to be more 'studenty' when I'm not in hospital and I'm supposed to more 'medically' when I'm in hospital. I kind of split it like that.*"

Virtual identity negotiation

Students spoke about a separation of their personal and professional identity on Facebook:

"*Like keeping it completely separate. … I have a lot of people I know in the course I don't have very many medical friends on Facebook because I want to keep it completely separate and people can't find me because … I know that can affect your career later, so part of me wants to quit [Facebook] anyway."*

Switching it on

A major theme to emerge was students' ability to switch into a professional mode on demand. The subtheme 'dressing up' was associated with this.

Participants discounted the applicability of professional standards outside of the clinical setting and talk of "*acting professionally*" and the ability to "switch it on" within the clinical environment.

"*… we're probably all acting professionally in the situation we're in, as students, we're acting as we should as students, … when you go into the hospitals you have to act differently around patients …*"

Dressing up

'Dressing up' as a doctor was a strong element of students' perceptions of being a professional. Clothing appears to be a means by which students "switch on" their professionalism.

"*… you dress smart, people will see you as professional…*"

"...we all dress up in our smart gear and ...it kind of gets you into a different mind set..."

Leniency

The penultimate major theme was leniency. Students strongly call for leniency regarding professional behaviour in Phase 1. The developing nature of professionalism was a subtheme.

"*I think we're allowed to be a little bit unprofessional, it's sort of student life isn't it? It's the two years of medicine that you actually get to be a dosser*."*

*Dosser: British slang for a lazy person or idler.

"I don't think we should be expected the same amount of professionalism as proper doctors because we are learning, aren't we?"

Developing professionals

Calls for leniency were accompanied by descriptions of professionalism as a developing, scalar attribute.

“We are learning also to be professional within the next five years and we will arise to it within the next five years and we will get better and better at it through the course.”

Sacrifice

The final major theme to emerge from the data were sacrifice. Students’ narratives described a feeling of being constantly observed and likened this to making a sacrifice regarding their freedom as an individual.

“There is a certain amount of freedom that people need to have though, … how much of your freedom is it right to sacrifice for a career? As a person, as an individual? … It can be tough to understand that that’s what you’re giving up to be a doctor. And you are giving it up.”

8.10 Discussion

Of the three contexts, university, virtual and clinical, undergraduate medical students feel that it is the clinical context in which professionalism is most relevant. This finding is not surprising, as it is often taught within this context by practising clinicians. Students’ professional development is facilitated by both clinical and academic staff; however, students appear to place more significance on the clinical perspective.

Lave and Wenger [291] coined the term “legitimate peripheral participation” as a means of describing the view that learning is a situated activity, in which the learner ultimately becomes a participant in the community of both knowledge and practice. This occurs when the learner masters skills and knowledge and becomes embedded within the sociocultural practices of the community. Eraut describes this as ‘learning from experience,”[292] and also attributes learning to the social interactions within the place of work.

As undergraduates, medical students are at the periphery of their eventual community of practice, the health care system. Not yet fully skilled, they acquire knowledge from the masters and develop their attitudes by observing the community. This development of sociocultural attitudes and skills, whilst becoming integrated into the community, may occur on placements and from observing role models in both the university and clinical context.

Learning, as it normally occurs, is a function of the activity, culture and content in which it occurs (the community) [291]. This is also true of professionalism, despite the lack of consensus on whether professionalism can be learnt; professional behaviours and attitudes are developed in the community of practice. Students’ future clinical practice is moulded through direct observation of experts, those whom students perceive to be the masters. Learning has been described as context dependent[291, 292] . This parallels the testimonies of the students that professionalism is best learnt within the practice community, and specifically through observation of role models.

Students described role modelling in the clinical context explicitly, both anecdotally and as a favoured means by which professionalism can be taught.

Role models were on occasions described positively, however most students gave accounts of lapses of professionalism and associated “camaraderie”. This behaviour was sign-posted by “role models” as acceptable, with students drawing the conclusion that it is legitimate to behave in an unprofessional manner provided that patients do not witness it. Narratives of events where patients had been mocked by role models were provided which were consistent with issues in current literature around professional socialisation and the use of cynical humour about patients[293]. Students seemed to find the associated camaraderie with their superiors as gratifying and a sign of acceptance into “cultural insider-knowledge”[294]. These findings corroborate recent research reporting that students experience pressure to conform to and adopt values and behaviours not acknowledged by the formal curriculum [295].

The “two gears” of professionalism, professionalism displayed in the presence of the patient and that displayed amongst peers, were strongly advocated by students as they had witnessed their superiors doing so. Furthermore, the “banter*” (*British slang for the exchange of light or teasing remarks; good-natured raillery) described was condoned under the premise that camaraderie and banter are acceptable methods by which medics can unwind. Perhaps a representation of Lave and Wenger’s statement that an understanding of “what old-timers collaborate, collude, and collide, and what they enjoy, dislike, respect and admire”[291] is required to become a full practitioner. Similarly, this mirrors the findings of Turner[296] that within new groups people surrender their own idiosyncratic views in

favour of those supported by the group. In other words, the students may not truly condone the behaviour of their role models but they accept it.

Literature suggests that “role modelling is fundamental to developing professional behaviour”[297], with some studies suggesting they are the primary influence[275]. With this in mind, educators, especially within the clinical context, must be aware that students appear to be acquiring professional behaviours by imitation of those they perceive to be role models. In order to uphold good professional standards and negate effects of negative role modelling, great care must be taken to sign-post appropriate professional behaviours, thus ensuring that students eventually become legitimate members of the clinical community. This is in accordance with suggestions by Baernstein et al (2009) that, “medical schools should ensure that students are exposed to excellent role models - ideally, faculty who can articulate the ideals of professionalism and work with students longitudinally in a clinical setting.”[275]

Identity was a recurring theme within the data. Identity represents the process by which the person seeks to integrate their various statuses and roles, as well as their diverse experiences, into a coherent image of self[298]. Epstein’s statement highlights a complex phenomenon with which most individuals struggle, “Who am I?” This question is likely to be asked by many medical students, both in their capacity as a learner and an individual. Students’ statements were indicative of personal grappling with identity and social standing, often implicitly resentful of such a dilemma. Whilst students acknowledge an understanding that their behaviour will be under scrutiny throughout their careers, it was accompanied by a sense of loss of identity. Students often made a clear distinction between themselves as a

doctor and themselves as individual, but expressed concerns that society and their community of practice failed to do so. However, literature reports medical students' identities being closely tied to their chosen profession, "often leading to deep internal prohibition of criticism of the profession or of fellow physicians."[299]

This dilemma draws a parallel with the notion of identity negotiation, whereby an individual negotiates with society regarding the meaning of their identity[300]. Identity negotiation in this group of students relates to the context and applicability of professionalism, as identified by themselves, in that students are confused as to when and where professionalism is required, and to what degree. As a student should the same professional standards be required as is required of a practising physician? Is professionalism required out of hours? Am I an individual or always a professional; should there be a distinction?

Identity negotiation is analogous with professional socialisation. Professional socialisation, the process by which individuals learns both the roles and responsibilities of their chosen profession before emerging as a member of the professional culture, is also comparable to the mastery of a profession as described by Lave and Wegner [291].

Professional socialisation is tri-phasic: recruitment, professional preparation and organisational socialisation. Recruitment and professional preparation are defined as the pre-service phases. These anticipatory phases occur before and during the professional education period [301]. Klossner suggests that there is overlap between the three phases. Furthermore, legitimation was found to initiate the process of professional socialisation. This

was due to its role in stimulating “meaningful experiential learning and, thus, formation of professional identity.”[301] When students look to others for acceptance, in order to affirm their developing professional identity, legitimation occurs [301].

By considering legitimation, situated learning, professional socialisation and identity negotiation, the comments of our students regarding camaraderie and role modelling are contextualised and perhaps comprehensible. In order to develop a professional identity, students must gain a realistic view of the challenges, and opportunities, of a profession[302]. Experiencing unprofessional camaraderie and role modelling, followed by negotiating a moral judgement on whether to engage with it, may be a learning opportunity for students. The learning opportunity presented will challenge students with respect to professionalism and professional identity.

The conflict between self and profession continues in the virtual environment. To students, both online and offline environments are their realities, these environments are equally constitutive. Online and offline life will continue to exist side by side[303]; professionalism is applicable to both.

In our data, responses regarding the virtual context of professionalism, specifically views on social networking sites and the rights of a medical student as an individual, were opposing. Most examples made reference to the social networking site ‘Facebook’. Examples were given of peers posting details of the course and clinical experiences as their status on Facebook. Some students made a clear distinction between their working and personal

lives, and could not understand why colleagues posted this type of information onto public forums. This once again reinforced students' struggle with identity negotiation, but from the perspective of another peer's identity. Students describe airing caution and the use of privacy settings on Facebook, or fostering a pseudonym so they could not be identified. There was awareness that lecturers and tutors used the site, so students could be accountable for any issues resulting from online social networking activity. Recent research on this topic[304, 305] also reported concerns with regard to inappropriate and unprofessional behaviours being publicly accessible on individual profiles and in groups.

Literature on gaming sheds light on the virtual world, to which this generation of students are well accustomed. A "virtual mask", in the form of an avatar is described, whereby the avatar provides opportunities for the player to experiment with identity in original ways. Adoption of such a mask secures both anonymity and identity simultaneously; this enables the player to be both known and unknown concurrently [303]. This gaming behaviour resonates with the behaviour displayed by the medical students on social networking sites such as Facebook, for example the acquisition of a pseudonym as described previously. Heim states that a virtual existence allows people to keep a distance while at the same time putting themselves on the line[306]. This appears to be the case for medical students; they feel able to express unprofessional attitudes on social networking sites by hiding behind their virtual profile, which may be reality, an idealised, or a masked version of themselves. This online, masked persona enables students to distance themselves from offline reality, the clinical community, and express attitudes that they may not otherwise feel comfortable projecting.

Incidents of negative role modelling online were reported demonstrating the extent to which social networking influences students and their professional practice. As social networking continues to define social progress, this raises questions with regard to professionalism and the public domain, something individuals and institutions may wish to consider. Institutions may need to formalise their stance on behaviour exhibited online, as we feel that this is likely to be an issue raised in future fitness to practice proceedings at some institutions. Crowe states that, “Cyber space, and its practices, clearly have potential for fomenting new moral panics centred on young people’s leisure lives.”[303] The Times reported the suspension of a Swedish nurse who had posted photographs of surgical procedures on Facebook [307] in 2008. The article suggests that posting of inappropriate material, such as photographs from clinics, is “designed to provoke envy among friends.”[307] Institutions are now faced with a dilemma, do they hold students accountable on the grounds of Fitness to Practice over violations of privacy rules and other related unprofessional postings on social networking sites? Could social networking be just another moral panic like that described by Crowe?[303] Do medical schools have the right to scrutinise students’ profiles, whether they be private or publicly accessible?

There was an overwhelming call from participants for institutions to be lenient with regard to professional standards whilst students are in an undergraduate setting. The rationale for this was multi-factorial including age, hierarchical status and inexperience amongst others. Coupled with this was the notion of a developing professional; students often described the construct of professionalism as a cumulative attribute which culminates at the end of the 5 year undergraduate period when students graduate as a professional, or “proper doctor”.

These comments are indicative of “proto-professionalism” as described by Hilton and Slotnick[278] which encompasses the developmental period of professionalism, prior to the students having clinical experience. Without this practical experience the evolution of students’ knowledge and skills base is limited. During this ‘proto-professionalism’ phase their behaviour is perhaps determined by moral and reflective judgements rather than by their experiential learning and subsequent practical wisdom[278]. Conversely, Cordingley *et al* [308] demonstrated that as students progress through medical school their levels of moral reasoning and ethical sensitivity do not increase.

Students’ descriptions were of professionalism as a persona and a professional as a character that they are able to act on demand. This was further supported by their accounts of dressing as a professional. When discussing professional behaviour and attributes, there was a surprising emphasis upon clothing. Students’ remarks indicated an element of costume dressing; they dress-up as they perceive a doctor should. Attire therefore made them a professional. Their struggle with identity and social standing becomes evident once again. Students are mimicking, through dress, the masters of their community of practice in order to become legitimate participators, “How masters talk, walk, work, and generally conduct their lives.”[291] The emphasis on clothing ought not to be dismissed. Research has shown that simulating seemingly trivial aspects of the clinical context, such as clothing, in undergraduate teaching demonstrates that for a high degree of authenticity in simulation approaches, small details are important[309, 310].

To conclude, the most consistent theme in our data were that of the context of professionalism; in which context is it relevant and to what extent, and in which context is a

student a medic or an individual? There was a tremendous sense of confusion evident across student bodies at both institutions. More striking was the impact role models were having on undergraduate students, and the magnitude of the less than positive hidden curriculum. We have demonstrated that students were aware that professional behaviour is expected of them; however the level to which it should be exhibited is not clear. Interestingly, professional behaviour was described as a "burden" and similies such as "being like a robot" were used. Upholding professional standards was perceived to be sacrificing the freedom of the individual; this was mostly attributed to "being watched". This was once more suggestive of the students' struggles with personal and professional identity and the influence of societal expectations upon this. More importantly, this should highlight to faculty the need for explicit guidance on professional standards and expectations of students as both trainee medics and as individuals. Furthermore, there should be awareness that this scrutiny is not always welcome and may be having a detrimental effect upon students. A "*Student BMJ"* article [311] reports arguments against measuring professionalism; "overly intrusive assessment" of professionalism was reported to be a cause of concern and anger for medical students [311, 312].

Lessons can be learnt from both parts of this study. These are that: students' acceptability is an important consideration when designing a peer assessment exercise, and that medical students remain confused by expectations of them with regard to professional behaviour. Faculty must sign-post their expectations clearly, and ensure that these expectations are reinforced by positive role-modelling.

8.10.1 Limitations and future work

The limitations of the study include the voluntary nature of participation, producing a biased population, and focus group dynamics. Focus groups were utilised because they allowed the researchers access to multiple, diverse opinions in an efficient manner. The information exchange within the group can be interactive and dynamic, allowing exploration of contrary opinions and reflection. Authors acknowledge that focus group dynamics can be easily changed by one individual who may be influential or opinionated; this limitation is inherent to focus groups. The outspoken participants may silence any contrary views; Noelle-Neumann (1984) described this as the Spiral of Silence Theory whereby people shape their opinions to prevailing attitudes about what is acceptable[164]. Furthermore, the presence of other research participants may compromise the confidentiality of the focus group session[313]. However, this was kept to a minimum by using experienced facilitators and establishing ground rules at the beginning of all focus groups. Ground rules were established by the facilitators and addressed issues of confidentiality, speaking in turn and accepting others' views. These rules enabled participants to feel comfortable discussing personal experiences.

A further limitation of qualitative research is potential author and facilitator bias. Within focus groups, facilitators may have presuppositions. By adhering to the focus group spines and by choosing facilitators who have no role in the teaching of professionalism we have minimised this. With regard to author bias, coding data independently initially, followed by collective axial and selective coding, minimises the personal interpretation of data from authors. Further research on the themes identified in this paper are required to further explore students' understanding of the development of their own professional behaviours.

Chapter 9: Exploring reflective 'critical incident' documentation of professionalism lapses in a medical undergraduate setting

9.1 Background

9.1.1 Contributions to this chapter

This study has been published in the journal BMC Medical Education (online)[135]. I was responsible for data analysis and was heavily involved in the preparation of the manuscript published.

Professor John McLachlan designed this study, with input from myself and Dr David Hodges. Data were routinely collected. Dr Hodges was also involved in data collection and manuscript preparation. All authors analysed the data, with negotiation.

9.2 Introduction

Measuring professionalism in undergraduate medical students is a difficult process, and no one method has currently emerged as the definitive means of assessment in this field [314]. Recent literature reviews have highlighted the lack of current evidence when measuring attitudes related to professionalism in medical education[271] and have proposed 88 methods in the past 25 years[152]. This study examines critical incident reporting combined with voluntary student reflection within a UK medical school as a method for exploring

professional behaviours and attitudes in a cohort of pre-clinical undergraduate medical students.

The advantages of critical incident reporting were first published by Rhoton, documenting the study of anaesthesiology residents in a US teaching hospital during a 5 year period in the 1980s[133, 134]. These studies found that only 1% of negative behaviour related to areas of professionalism, and that there was a link between resident conscientiousness and unprofessional behaviour. Further to this, critical incident reporting was found to permit early intervention in unacceptable behaviour [132]. Incident reporting forms in recent studies demonstrate the use of areas of fixed concern, derived from professional standards the faculty feel students should exhibit, as a method to assess professionalism [315, 316]. Critical incident reporting can be used as grounds for dismissal if continuing unprofessional behaviour is observed[315]. In the University of California, San Francisco, critical incident reports formed part of the medical student performance evaluation given on graduation if two or more reports had been issued during undergraduate study [315]. Outside of medical education, critical incident reporting remains an important aspect of improving patient safety and is used in all NHS trusts to monitor adverse events[317].

Student reflective skills have been seen to be highly important in the development of professional behaviours [318]. By studying student reflections on critical incident behaviour, the development of good professional behaviours from reflection could be explored. Its introduction could provide tools for students to examine their own behaviour and attitudes, to consider the reasons why incidents occurred, and to allow solutions to be developed for preventing such incidents in the future.

9.3 Methods

Existing literature was reviewed to look for current evidence in this field. Medline was searched using keywords of “professionalism”, “critical incident”, “incident reporting”, “medical school” and “undergraduate”. Further manual searches within three leading journals (*Medical Teacher, Medical Education and Academic Medicine*) were made to ensure all relevant papers were considered.

This study used an unrestrictive method of incident reporting to record unprofessional behaviour utilising expert opinion by a process of connoisseurship[319]. Experienced staff members reviewed all incident reports and made decisions based on knowledge and personal experience as to whether the reports were valid and primarily focused on unprofessional behaviour. Unprofessional behaviour was defined by any act which broke the School code of conduct, which is based on the GMC *Duties of a Doctor* [20, 103, 230]. Students were given a copy of the report prior to a meeting with staff to discuss it, and invited to submit a reflective response using a structured form.

The study group was 228 undergraduate medical students in Years 1 and 2 of a U.K. medical school. The study material was all critical incident reports issued during the timeframe of a single academic year. The exclusion criteria were reports that did not refer to professional behaviour and incomplete reports. Framework analysis was done using a process of thematic analysis, and the data collected were analysed using a grounded theory approach[158]. All analysis took place at the end of the study period. Initial multiple

readings of the transcripts with hand annotation was performed, before incorporating the data into word processor format for further study. All three authors analysed the data independently, with comparative discussion of findings performed to reach consensus and promote reliability of results.

The data used in this study were obtained from routinely collected information regarding student assessment and was not initially intended for the purpose of research. Subsequently, the information obtained proved, we believe, to be of generalisable interest [320]. Ethical approval from the University Ethics Committee for the use of educational research data were obtained, on condition that the data were anonymous and could not be linked to the individuals concerned. As a result, individual student consent was not obtained. All student data were anonymised, identifiable comments were excluded, and no identifiable harms arose from this study being conducted.

9.4 Results

Over the study period, 16 critical incident forms were completed by members of faculty from a cohort of 228 students. Of these 2 were excluded following connoisseur review as they failed to meet the criteria set. The 14 forms related to 9 students, some of whom had multiple forms. Of the 14 incidents forms, 7 had a reflective response.

Seven elements of professionalism were raised by staff. These were poor communication, unexplained absence, record keeping, meeting deadlines, positive commitment to studying,

honesty and patient confidentiality. Students did not always recognise the same areas of professionalism as a concern within their reflections.

The following major themes emerged from analysis of the transcripts.

Impacts and consequences

All students who reflected on the incidents considered the impact and consequences of their actions. There was recognition of how the student was involved.

"This affected me and my studies."

There was also recognition that members of staff, patients and other students may also have been affected. Students described the impact that their actions had on others, what burdens may have been placed upon them and what additional work, if any, the student had created for them.

"It affected me and the people at my placement. I was not as useful and productive as I could have been…it affected my colleagues at the placement as they may have had to perform a greater share of the work that would have otherwise been the case."

"I did not attend my sessions in which I appreciate my clinical tutor has given up valuable time to come in."

The ideas of wasting the time of others and failure to complete their own work emerge from the explanations. Students are also concerned about how the incident may have changed perceptions of themselves and other students.

"...my patient may have been let down and lost respect for the responsibility of medical students."

The concern of student concentrates on patient perception of the student which has been affected in a negative manner.

Reflections

Identification of the responsibilities of a medical student is a necessary step to remediation. These reflections on the incidents fell into two sub themes: those of acceptance and those of rejection.

Acceptance Reactions

With acceptance reactions, students demonstrated the recognition of fault in themselves, showed guilt and remorse for their actions and acknowledged a lack of knowledge could be the cause. Students included clear statements of acceptance.

“I fully accept that my behaviour has been unacceptable.”

“My failure to provide explanations and consult with the tutor is unacceptable.”

Students also show the ability to describe how the event has affected them emotionally and their intended future response.

“I feel very bad that I let this happen and it certainly won’t happen again.”

“I now realise that it was very serious and should not even have considered it.”

There is also student acceptance of blame which had been previously unrecognised.

“*For some reason it did not occur to me that the same principles applied in the lecture theatre.”*

Rejection reactions

Not all students in the study felt the reports were an accurate description of the event which occurred, and described aspects which they felt were out of their control.

“*…haven’t received the test results from the medical centre.”*

“I was falsely led to believe…”

Denial of the accuracy of the event also occurred, and students disputed the incident. Students felt that their actions have been misinterpreted by the staff and that their actions were not unprofessional. These rejection reactions could also be seen to be both defiant and possibly complacent.

“*I was not listening to music during lectures.”*

Some responses do not deny that the event occurred, but reflected a denial of the significance of the event which occurred. Students did not always agree that what was reported constituted a critical incident or problem with professionalism.

“*No event, I missed a few important lectures*.”

A minority of students recognised that their behaviour was unprofessional from the view of staff involved, but believed that their actions were justified due to the circumstances surrounding the issue.

“I worked most of the holiday and believe this time would have been necessary regardless.”

"...due to lack of work earlier in the year that I needed to catch up."

"… but I believe that the alternative would have hindered my revision."

Responses

Students frequently identified further actions, which in turn related to whether they had made acceptance or denial responses. Acceptance responses included remediation in the form of apologies:

"*I will write a personal letter of apology."*

"*I take full responsibility for my actions and sincerely apologise to both X and Y, whose lectures I enjoy very much."*

Apologies can also relate to the theme of impact and consequences, with students being concerned with the perceptions staff may have of them in light of the incident. By apologising, students may hope to prevent this perception of them.

Rejection responses include suggestions for further actions by others, which could be constructive. Students feel that their experiences could be used to make practical changes to the course, and would as such be of benefit to their fellow students.

"...instead of having three lectures on one topic, the student should be given a book to study with a certain time period. Then at the end of this time period there will be a review lecture to ensure everyone has understood."

9.5 Discussion

The approach of having critical incident reporting forms associated with reflective documents appears to have the possibility of promoting student reflection as a necessary first step to remediation of inappropriate behaviours. Suggesting practical steps to avoid the repeat of the situation in the future shows potential progression through the reflective cycle, with the hope that future behaviour will follow the changes the student suggests they could make. Even having such a system in place may serve as a promoter of good behaviour or an inhibitor of adverse behaviour, since students know that the critical incident form will be placed in their file. The evidence of Papadakis et al. suggests that such incidents can serve as a predictor of future behaviour, and an organised system of collection of such data could facilitate future progression decisions, or simplify the task of providing references or Dean's Letters[109, 110, 236, 315, 321].

The biggest barrier to gaining a positive outcome for faculty and student occurred when no reflective response was written by the student. In future use of the critical incident reporting, making reflection a compulsory aspect of the incident becomes an attractive proposition. An absence of reflection only has a negative value in assessing the student's professional behaviour, whereas student reflection can have either positive or negative value.

The curriculum of medical schools already incorporates reflection as part of summative assessment for many different taught subjects, and the “Report and Reflect” system could be used in a similar manner. Establishing expected outcomes for reflection and having procedures in place for students who fail to meet the outcomes would allow for a structured assessment process for students involved in critical incidents. As a summative assessment it would also have formative value, the student would learn what is expected in future as a result, and would be able to change their professional behaviour as a result.

Although the sample size in this study was small categorisation of Critical Incidents could be used to shed light on gaps in the professionalism curriculum, as has been demonstrated in other studies[322, 323]. Ginsburg and colleagues categorised Critical Incident reports from three universities. Most incidents were witnessed or known about, as opposed to self-reported. They found that six critical issues emerged. These were “communicative violations (to or about patients or other health care professionals); role resistance (individuals chafing against constraints or expectations of their perceived roles); objectification of patients (ignoring patients or treating patients as vehicles for learning); accountability (to colleagues or patients, including avoiding patients, failing to disclose information, or failing to treat appropriately); physical harm (to patients or others); and crossfire (being put in the middle of a struggle between superiors).”[323] The authors concluded that the critical issues reported did not map easily onto standard, abstract definitions of professionalism. Furthermore, they indicated that this “incongruence suggested that the development of effective curricula in this domain must bridge the gap between traditional taxonomies and students' perceptions of professionalism”.[323]

By collecting further Critical Incident evidence, faculties may be able to identify gaps in their curriculum, and as in this chapter, reflections on critical incidents may also provide useful information on the areas of professional conduct that students are struggling to either understand or adhere to. In a further study, Ginsburg and colleagues suggested that faculty should highlight the similarities between medicine and other professions as a means of teaching about professionalism, as opposed to stressing the differences[322]. “As educators, rather than beginning medical students’ professionalisation by drawing attention to what makes them different from others, perhaps we should reinforce what they have in common with other professions by ingraining them with a sense of citizenship in the broad community of professionals before emphasising those few but critical principles that set medicine apart from other professions. If we were to encourage an appreciation of generic professional values first, and then to layer on specific values such as altruism, perhaps we would have more success in cultivating professionalism without entitlement, altruism without self-neglect and expertise without omniscience (p18).” Although Critical Incidents are useful as a means by which unprofessional behaviour of a single student can be documented, the comments of Ginsburg and colleagues [322] demonstrate how such reports can be utilised for development of all students, with specific reference to the professionalism curriculum.

With regard to this study, I acknowledge that there are limitations to the approach used. No evidence of data saturation was observed in the responses, and the study needs to be extended. Since critical incidents are rare events, extended periods of time will be required to build up a full analysis of causes and responses. The cohort of students involved in this study was small at 228, and the timeframe of single academic year was short. Involving multiple centres and increasing the timescale to the extent of 5-10 years, with further follow-

up studies of students once working as doctors, would allow for improved reliability and could give evidence of predictive validity.

The utility of this approach can also be considered [150]. Reliability is dependent upon having staff with sufficient knowledge and experience to assess the reports. The educational impact should be positive, which can be seen when reflective practice improves future professional behaviour. Staff acceptability may relate to the time involved, with each report taking on average 90 minutes to complete in this study. Issues of validity relate to the GMC requirements for student behaviour, and for medical school monitoring of such behaviour.

9.5.1 Conclusions

This study shows that by using a “Report and Reflect” system of incident reporting, a strong platform can be created from which lapses in professional behaviour within the student body and be recorded and challenged. It has been seen that student reflection on critical incidents could promote positive changes in professional behaviour. Further to this, an absence of reflection, or the inability of a student to respond adequately to events, might be a future indicator of unprofessional behaviour. Critical incident reports are enhanced when combined with student reflection.

9.5.2 Future work

This study could be repeated with subsequent cohorts in order to determine whether any patterns of behaviour emerge. A cohort study could be carried out once students graduate

to correlate critical incidents recorded as an undergraduate with postgraduate fitness to practice proceedings.

Chapter 10: Discussion

As aforementioned in the introduction, this thesis describes two aspects of undergraduate medical training: anatomy teaching and professionalism. For the purposes of this discussion I will address each in turn, highlighting the commonalities to emerge from my two areas of research. I will also describe my findings in light of educational theory.

I have discussed each study in this thesis within the appropriate chapter. This discussion serves to highlight the common themes which have emerged from the research and how the phenomena described could be investigated further. The commonalities described later in this discussion demonstrate how the two distinct tracks of research undertaken, have partially united.

10.1 Innovations in anatomy teaching

The latest GMC guidelines (Tomorrow's Doctors, 2009)[21] call for medical schools to consider providing initial training in a clinical skills facility to minimise the risk to patients, ".... students should have the opportunity to become increasingly competent in their clinical skills..." The approaches to anatomy teaching discussed in this thesis, including the use of the Virtual Human Dissector, the use of body paints and simulation using scrubs, enable us to work towards that goal. Budget restrictions hinder the introduction of high fidelity simulation suites across many UK medical schools. However, I have demonstrated how the use of the aforementioned interventions can set anatomy teaching within a more clinical context and can enable the early introduction of clinical skills through using peer contact and have highlighted how the use of simulated patients (life models) could also be used,

thus simulating clinical practice. This clinical emphasis, as called for by the GMC is likely to have direct, positive outcomes with regard to patient safety and care.

Medical students require a good scientific grounding but must be competent to practice (to 'do'). When designing the medical curriculum, or implementing an innovation within it, one must always consider this course aim. It is my opinion that the purpose of the undergraduate curriculum is to produce Foundation doctors who are safe to practice, not specialists. The undergraduate course, whilst important, is essentially a generic training programme; more detailed anatomical knowledge is acquired during postgraduate training.

The aim of any clinical skills curriculum is to teach an explicit set of clinical competencies which must be mastered by medical students. I believe that this can only truly be achieved by integration between anatomy and clinical skills curricula. Skills must be linked to structure and function, grounded in basic sciences, with a focus on the living body. This should facilitate transferability from the educational to the practical (clinical) setting.

The purpose of the anatomy curriculum is therefore to provide the scientific background to enable clinical practice. As educators we must remember that we are creating doctors, not anatomists. Anatomical information is clinically relevant and it therefore is essential for the acquisition and retention of clinical knowledge and skills.

In Chapter 4, I suggested the notion of “learning landmarks”. These vivid experiences, which are memorable in themselves, provide access to the educational content associated with that context[90]. I have described how anatomy teaching can create learning landmarks/teaching tools, thus making the content more memorable. These versatile teaching tools promote students to relate the anatomy learnt to the clinical skills which are often taught concurrently. Learning landmarks enable emotional engagement (for example by being examined or painted), thus making content easier to recall. It is my firm belief that effective curriculum development is outcome driven. In the undergraduate setting the outcome for learning anatomy is to enable clinical practice, and thus creating such landmarks integrates the desired outcome with the intermediary steps (undergraduate education).

Clinical skills are reinforced by combining visualisation of anatomical structures, such as body painting and use of the VHD, with simultaneous examination. These methodologies keep the anatomy, the skill and the patient experience as the foci. These methodologies create learning landmarks, which in turn are likely to promote deep learning, primarily through emotional engagement in the learning activity. One conclusion which can be drawn from the data presented in this thesis is that experiential learning is powerful, as shown by the data presented on body painting and the use of clothing in simulation.

Another take home message is the importance and value of simulation in medical education. I have demonstrated how content is best remembered in the same context as that in which it was learnt. This was achieved by simulating the practice environment (clinic) in the educational environment (anatomy lab), the headline being that small details, such as

clothing, matter when simulating the clinical context. While the context of this intervention was recall of anatomical information, simulation has direct implications on patient care. Simulation provides a safe environment for the learner, amongst many other things, to make mistakes. It is my hope that students will remember the anatomy they have previously learnt when examining patients on the ward if we mimic the clinical context in undergraduate teaching, that is, simulation of the clinical context in the education context will promote future recall of anatomy and associated skills.

Best practice in anatomy teaching has been culturally created. For centuries we have been socialised into using cadaveric specimens. This is an expected tradition and for some, a rite of passage. Much of the research in anatomy education has focussed on the use of dissection as a teaching tool. The educational value of dissection remains unresolved since there are no controls comparing the clinically relevant outcomes of dissection based and non-cadaveric anatomy courses. I hope that the anatomical pedagogy research disseminated from this thesis[69, 87, 90, 187] will encourage anatomy teaching which is not only informative and explanatory within the scientific paradigm, but promotes students to consider their future patients. This could be achieved by signposting the clinical relevance of the content taught and simulating the clinical context but also by promoting the use of teaching methodologies which stimulate students and provoke narratives. It is only when these goals are achieved that we take steps towards reversing the desensitisation of medical students which has long been viewed as a potential outcome of medical school.

Implementation of the initiatives in this thesis are not without their risk. One could envisage how without careful planning and embedding into the curriculum an educational innovation

would fail. Educators should strive to balance the risk of fragmentation of the curriculum and potential loss of integration with innovation and forward thinking. Further balance must be achieved between transformational change, sustainability and the principles of good educational practice. The most important considerations must remain; what are the benefits to the students and how will this improve patient safety and care? Anatomists and clinicians should only observe anatomy tradition if it is evidenced good practice in these regards.

With regard to the innovations described in this thesis, sustainability must be considered. Any introduction into the curriculum must be cost effective. If an intervention cannot be rolled out across cohorts then its educational impact is diminished.

The work presented in this thesis has already begun to influence the practice of others. The qualitative data presented on students' attitudes towards body painting has been cited[19]. Sugand *et al* (2010) have used the qualitative data published[90] to advocate to others the usefulness of using body painting as a medium by which greater body comfort can be gained, as well as supporting the evidence presented by myself [90] and others[88, 89] that body painting is a valuable resource for teaching surface and living anatomy.

This thesis has identified factors such as context (scrubs) and colour (body painting) which have been measured and reported, respectively, as impacting upon retention of knowledge, specifically with relation to anatomy pedagogy.

It has been said that “medical education should be informed by a research-based theory of how people learn (i.e. the science of learning) and evidence-based principles for how to design effective instruction (i.e. the science of instruction).”[324]

Medical education is described as a field which often involves multimedia learning[324]. Mayer defines multimedia learning as from words and pictures, which may require the combination of both verbal and pictorial learning.

“Verbal learning involves learning with printed words (such as bullet points in a slide presentation or words printed in a textbook or on-screen text in a computer-based lesson) or spoken words (such as the speaker’s voice in a slide presentation or the narrator’s voice in a computer-based lesson). Pictorial learning involves learning with static graphics (such as illustrations, diagrams, photographs, drawings or charts) or dynamic graphics (such as animation or video).”[324]

Mayer described the active processing principle of learning which proposes that meaningful learning occurs when learners engage in appropriate cognitive processing during learning, including attending to relevant material, mentally organising it into a coherent cognitive representation, and integrating it with prior knowledge activated from long-term memory[324]. It is my opinion that the theory of active processing has commonalities with principles of surface and deep learning as presented by Marton and Säljö (1976)[325].

Marton and Säljö[325] derived two approaches to study, “surface” and “deep” learning, as a result of empirical research, which have since elaborated by many, including

Ramsden[326], Biggs[327, 328] and Entwistle[329]. The deep and surface learning styles are best compared directly, as shown in table 18:

Surface	Deep
Focus is on the 'signs' (or on the learning as a signifier of something else)	Focus is on 'what is signified'
Focus on unrelated parts of the task	Relates previous knowledge to new knowledge
Information for assessment is simply memorised	Relates knowledge from different courses
Facts and concepts are associated unreflectively	Relates theoretical ideas to everyday experience
Principles are not distinguished from examples	Relates and distinguishes evidence and argument
Task is treated as an external imposition	Organises and structures content into coherent whole
Emphasis is external, from demands of assessment	Emphasis is internal, from within the student

Table 18: A summary of deep and surface learning

(Based on Ramsden, 1988)

I have previously described surface and deep learning as part an assignment[330] for my Postgraduate Certificate in Teaching and Learning. It is important to draw the distinction between the individual and the learning approach. Despite learners being classified as "deep" or "surface", these are not attributes of individuals. Each is merely an approach to learning, and an individual may use both approaches at different times. Nonetheless learners display a preference for one or the other[331].

Learning style correlates with motivation. “Deep” learning correlates with intrinsic motivation and “surface” learning with extrinsic motivation, but they are not necessarily the same thing. Either approach can be adopted by a person with either motivation [330].

Students may vary in the way they approach learning for a specific topic. This is also true of how an individual learns different topics. Educators having an awareness of such variability in a learner’s approach will be enabled to adapt their teaching style to suit the needs of their learners. This reinforces the “student-centred” approach mentioned above. The innovations presented in this thesis work towards a student-centred approach to teaching. Furthermore, the learning landmarks which it is hoped the anatomy curriculum create may serve to engage students in deep learning approaches, rather than employing superficial rote learning to the acquisition of their anatomical knowledge.

Mayer describes the unlimited capacity of sensory and long-term memory for retaining information. However, working memory is described as possessing a limited capacity for processing information, something Mayer identifies as acting as a “bottleneck in the system.” The relevance to anatomy education, as well as education generally, is that with regard to working memory information needs condensing and organising into “meaningful chunks” in order for the learner to work within the constraints of their limited processing capacity. The psychology of the learning process is out with the remit of this thesis. However, the information presented by Mayer is a useful consideration when designing and implementing change into a medical curriculum, “People must be active learners who seek meaningful learning because they do not have the processing capacity to attend to every piece of information that is available to them.”

10.2 The Conscientiousness Index and professionalism

I have presented the Conscientiousness Index as a tool for assessing one potential element of professionalism, conscientiousness. The CI enables faculty to utilise routinely collected data in order to make objective, scalar measures of students' conscientiousness. I have shown how this measure of conscientiousness correlates with staff and peer judgements on professionalism. These data might suggest that conscientiousness is a proxy measure of a student's professionalism.

The Conscientiousness Index, peer assessment of professionalism and qualitative data presented in this thesis have served to highlight that professionalism remains a slippery concept. This is not only with regard to its definition but also its assessment. Peer assessment has raised concerns over students being asked to "rat on their mates" and the impact of collegiality on the validity of such assessments, while the Conscientiousness Index has received negative press in the national media[332], "Medics riled by tests of professional behaviour."

The unanswered question regarding the CI is that of its predictive validity. In order to ascertain whether the CI has predictive validity, a case control approach could be employed, since a cohort study is not practical at this moment in time, which would entail a retrospective inspection of the doctors' performance. Such a measure of their professional performance would be, for example, referral to the GMC for Fitness to Practice (FtP). Any

doctor with a FtP record who had been part of a cohort where CI data were collated could then be matched to 3 or 4 peers who did not have any FtP incidents on their record. These matches would be against demographics such as sex, entry status into medical school or ethnicity. The CI scores for these peers, plus that of the individual referred for FtP would then be compared.

For each case an increased number of control students would be compared as this would strengthen the statistics. It would obviously not be possible to increase the sample size of referrals as those individuals included would be limited to doctors who had both CI data and an FtP proceeding on their record. However, this case control approach would inevitably be limited by the Data Protection Act due to the ethical considerations for obtaining information and permissions from future employers regarding FtP files and staff records.

A desirable development of current UK practice would be for the commencement of a uniform tracking system for medics throughout their university and work based careers. It is believed that this is under consideration with the GMC, who, it is speculated, plan to register all undergraduate medical students. These plans were in response to a request from Judge Janet Smith, who chaired the Shipman Inquiry, that all medical students should be added to the GMC register[333]. Similarly, the United Kingdom Clinical Aptitude Test (UKCAT) Consortium Ltd[334], who conduct aptitude testing for all medical school entrants on behalf of 26 UK medical and dental schools, are proposing integrating their database of aptitude test results with the GMC to form a nationwide database[335]. Should the GMC plan to introduce revalidation for all doctors come to fruition, then this too will create a national database containing, among other things, information documenting concerns regarding

professional conduct. Any such database would be invaluable when assessing the predictive validity of the CI.

The quantitative approaches utilised in the analysis of CI data in this thesis were relatively unsophisticated. An investigation is currently underway, with the assistance of Dr Paul Tiffin, to explore the relationship between performance on exam questions labelled as “professional behaviours questions”, “anatomy questions” and both CI and peer nomination data. Using Rasch analysis, a branch of Item Response Theory which draws interesting parallels with multi-level modelling, Professor McLachlan, Dr Tiffin and myself are working towards to exploring the efficacy of professional behaviours questions, which have long been relied upon as a measure of professionalism. Such questions are often weak predictors and, we believe, have no obvious value.

Dissemination of research regarding the CI [92, 93, 223, 224, 226-229, 336, 337] has yielded the employment of the index across medical discipline such as general practice, surgery and anaesthesia, as well as its uptake in professions including law and teaching. Time will tell what predictive validity the CI has in both medicine and other professions.

I have highlighted the diversity in views of professionalism and conscientiousness. Although definitions exist, I believe professionalism to be a moving target whereby the definition changes as societal and professional body expectations evolve. To address this uncertainty, a rolling consensus definition could be sought. This could be derived by employing a rolling Delphi study design to seek the input of medical educationalists,

practising clinicians and service users which might be students and patients. A period of five years could be sufficient to see a shift in expectations and thus the need for the next consensus. From the results of this thesis, it is clear that professionalism is context dependent. Governing bodies, such as the GMC, need to work towards producing clearer guidance for medics at each stage of their career. Progression of professional behaviours cannot be made if the individual concerned does not know what is expected of them.

The most recent context in which professionalism has become relevant is that of the internet. I envisage that much debate will persist in years to come regarding the applicability of medical professionalism within virtual contexts, such as email, websites, chat rooms and social networking sites. This thesis has served to highlight a student perceived lack of clarity on the behalf of faculties and governing bodies as to what is expected of medical professionals with respect to their conduct online.

10.3 Emergent themes

In addition to the emphasis placed on clothing in Chapter 3, students' attire emerged as a theme throughout this thesis. This result is of interest because within the professionalism context it emerged spontaneously. Students spoke of clothing as part of their professional identity. The role of clothing in professional identity formation is something that I would like to explore further in my future research. This research could include other professions, such as those that wear a uniform, as well as medicine and allied health disciplines.

Professional identity has also been described within the online and virtual context. The *virtual presence* of a student or clinician (i.e. their online activity), whether it is on social networking sites or in an email, is a reality to those viewing the content and to the individual. As proposed in Chapter 9, the GMC and individual institutions ought to guide medics as to expectations of professional conduct while utilising online or virtual environments, specifically when a medical student or clinician is outwith office hours.

10.4 Research paradigms

My thesis has employed methods from both the quantitative and qualitative paradigms. Debate regarding the efficacy of both approaches is long-standing. I have found that a relationship can be built between qualitative and quantitative data, which is helpful when answering a research question. The data sets are different yet interactive; each sheds light on the other. Within my results, qualitative data have aided my interpretation of phenomena which have been quantitatively derived.

10.5 Limitations

This thesis has undoubtedly been limited by the restricted cohort sizes at the medical school, as well as in most studies, the self-selecting bias of the participants. As described within individual chapters, where qualitative data has been presented I have endeavoured to reduce author bias and the effects of my presuppositions by utilising grounded theory and by coding with colleagues.

Another challenge that I have described is that of the real-world setting. Factors which have impacted upon the success of my research have included participant saturation and conducting research in a functioning department. Durham University's medical school is small, at any one time we have around 200 medical students. The department has an active Medical Education Research Group (MERG) which means that students often get saturated with requests for participation, often resulting in studies yielding low participation rates. Furthermore, students often tire of interventions, as in Chapter 5, as the educational setting is real-world. This means that as a researcher I had to respond to this feedback and change the study to a sub-optimal design.

Qualitative methodologies lent themselves to chapters in this thesis as they allowed for the discovery of participant opinions. Qualitative methods are not without their criticism. To increase the rigour in my data analysis I have been able to employ colleagues to code data. Respondent validation has also enhanced rigour.

All qualitative research interviews and focus groups undertaken utilised a semi-structured design. This approach does not impose pre-determined categorisation, therefore preventing the interview structure from limiting the field of inquiry. This flexible style enables participants to introduce their own agenda.

Generalisability and transferability from qualitative research have been cited as issues by many. Transferability implies that the onus is on the reader to judge the methods, context and results of a qualitative study and thus evaluate whether the findings are transferable to

their context. Where it was possible I enlisted participants from another medical school in order to highlight my findings as both generalisable and transferable.

10.6 Summary of future work

Qualitative research is planned to further explore professional and personal identities within undergraduate medical students.

A repeat of the body painting cross-over study is planned which will compare body painting and line drawing for sessions which require larger areas of block application to deduce the impact of colour on retention of knowledge.

Further critical incident data has been collected and will be compared to the conscientiousness index.

References

1. Dyer, G. and M. Thorndike, *Quidne Mortui Vivos Docent? The Evolving Purpose of Human Dissection in Medical Education.* Academic Medicine, 2000. **75**(10): p. 969-979.
2. Anatomical Society of Great Britain and Ireland, *A core syllabus in anatomy for medical students - Adding common sense to need to know.* Europen Journal of Anatomy, 2007. **11**(Supplement 1): p. 3-18.
3. Azer, S. and N. Eizenberg, *Do we need dissection in an integrated problem-based learning medical course? Perceptions of first- and second-year students.* Surgical and Radiolgical Anatomy 2007. **29**: p. 173-180.
4. Biasutto, S.N., L.I. Caussa, and L.E. Criado del Rı´o, *Teaching anatomy: Cadavers vs. computers?* Annals of Anatomy 2006. **188**: p. 187-190.
5. Drake, R.L., *Anatomy Education in a Changing Medical Curriculum.* The Anatomical Record (New Anat.), 1998. **253**(1): p. 28-31.
6. Gogalniceanu, P., et al., *Anatomy Teaching in the 21st Century—Dead Cool or Cold Dead?* Anatomical Sciences Education, 2008. **1**: p. 136-137.
7. Granger, N.A., *Dissection Laboratory is Vital to Medical Gross Anatomy Education.* The Anatomical Record (Part B: New Anatomy), 2004. **281B**: p. 6-8.
8. Unknown. *Ancient Egyptian Medicine*. [cited 30 October 2007]; Available from: (http://www.schoolshistory.org.uk/ancientegyptianmedicine.htm). .
9. Unknown. *Edwin Smith papyrus*. [cited 30 October 2007]; Available from: http://en.wikipedia.org/wiki/Edwin_Smith_papyrus).
10. Unknown. *Eyptian Medicine*. [cited 30 October 2007]; Available from: http://www.crystalinks.com/egyptmedicine.html.
11. McGrew, R.E., *Encyclopedia of Medical History*. 1985, New York: The Macmillan Press.
12. Porter, R., ed. *Cambridge Illustrated History: Medicine*. 2006, University Press: Cambridge.
13. Pranghoeffer, S., *The rete mirabile*, Centre for the History of Medicine, Durham University.
14. Unknown. *Barber-surgeons*. [cited 12 November 2007]; Available from: http://www.covent-garden.co.uk/historieso/trim.html.
15. Strathern, P., *A Brief History of Medicine: From Hippocrates to Gene Therapy*. 2005: Robinson.
16. Richardson, R., *Death, Dissection and the Destitute*. 2001, Chicago: University of Chicago Press.
17. Wikipedia. *Anatomy Act 1832*. [cited; Available from: http://en.wikipedia.org/wiki/Anatomy_Act.
18. Pabst, R., *Gross Anatomy: An Outdated Subject or an Essential Part of a Modern Medical Curriculum? Results of a questionnaire circulated to final-year medical students.* The Anatomical Record, 1993. **237**: p. 431-433.
19. Sugand, K., P. Abrahams, and A. Khurana, *The anatomy of anatomy: A review for its modernization* Anatomical Sciences Education, 2010. **3**(2): p. 83 - 93.
20. General Medical Council, *Tomorrow's Doctors*. 2003, General Medical Council.
21. General Medical Council, *Tomorrow's Doctors*. 2009, General Medical Council.

22. Heylings, D.J.A., *Anatomy 1999–2000: the curriculum, who teaches it and how?* Medical Education 2002. **36**: p. 702-710.
23. McLachlan, J., *New Path for Teaching Anatomy: Living Anatomy and Medical Imaging vs. Dissection.* The Anatomical Record (Part B: New Anat.), 2004. **281**(B): p. 4-5.
24. Parker, L.M., *Anatomical Dissection: Why are we cutting it out? Dissection in Undergraduate Teaching.* Australia and New Zealand Journal of Surgery 2002. **72**: p. 910-912.
25. Arráez-Aybar, L.-A., M.I. Casado-Morales, and G. Castaño-Collado, *Anxiety and dissection of the human cadaver: An unsolvable relationship?* The Anatomical Record Part B: The New Anatomist, 2004. **279B**(1): p. 16-23.
26. Arráez-Aybar, L.-A., G. Castaño-Collado, and M.-I. Casado-Morales, *Dissection as a modulator of emotional attitudes and reactions of future health professionals.* Medical Education, 2008. **42**(6): p. 563-571.
27. Bertman, S.L. and S.C.J. Marks, *The dissection experience as a laboratory for self-discovery about death and dying: Another side of clinical anatomy.* Clinical Anatomy, 1989. **2**(2): p. 103-113.
28. Granger, N.A. and D. Calleson, *The impact of alternating dissection on student performance in a medical anatomy course: Are dissection videos an effective substitute for actual dissection?* Clinical Anatomy, 2007. **20**(3): p. 315-321.
29. Lempp, H.K., *Perceptions of dissection by students in one medical school: beyond learning about anatomy. A qualitative study.* Medical Education, 2005. **39**: p. 318-325.
30. McGarvey, M.A., et al., *Dissection: A positive experience.* Clinical Anatomy, 2001. **14**(3): p. 227-230.
31. Schwartz, C.E., et al., *Impact of Cadaver Dissection: Working Toward Solutions.* Anatomical Sciences Education, 2008. **1**: p. 269.
32. Winkelmann, A., *Anatomical dissection as a teaching method in medical school: a review of the evidence.* Medical Education 2007. **41**: p. 15-22.
33. Winkelmann, A., S. Hendrix, and C. Kiessling, *What do students actually do during a dissection course? First steps towards understanding a complex learning, experience.* Academic Medicine, 2007. **82**(10): p. 989-995.
34. Yeager, V.L., *Learning Gross Anatomy:Dissection and prosection.* Clinical Anatomy, 1996. **9**: p. 57-59.
35. McLachlan, J.C., et al., *Teaching anatomy without cadavers.* Medical Education 2004. **38**: p. 418-424.
36. McLachlan, J.C. and S. Regan De Bere, *How We Teach Anatomy Without Cadavers.* The Clinical Teacher, 2004. **1**(2): p. 49-52.
37. Miles, K.A., *Diagnostic imaging in undergraduate medical education: an expanding role.* Clinical Radiology 2005. **60**: p. 742-745.
38. Cahill, D.R. and R.J. Leonard, *The Role of Computers and Dissection in Teaching Anatomy: A Comment.* Clinical Anatomy, 1997. **10**: p. 140-141.
39. McLachlan, J.C. and D. Patten, *Anatomy teaching: ghosts of the past, present and Future.* Medical Education 2006. **40**: p. 243-253.
40. Skidmore, J.R., *The Case for Prosection: Comment on R.L.M. Newell's Paper.* Clinical Anatomy, 1995. **8**: p. 128-130.
41. Malamed, S. and D. Seiden, *The Future of Gross Anatomy Teaching.* Clinical Anatomy 1995. **8** p. 294-296.
42. Raftery, A.T., *Anatomy teaching in the UK.* Surgery, 2007. **25**: p. 1-2.

43. Plack, M., *Computer-assisted instruction versus traditional instruction in teaching human gross anatomy.* Journal of Physical Therpay Education, 2000. **14**(1): p. 38-43.
44. Her Majesty's Inspector of Anatomy, *Human Tissue Act 2004- Explanatory notes*. 2004.
45. Her Majesty's Inspector of Anatomy, *Human Tissue Act 2004*. 2004.
46. Berube, D., *Cadaver and computer use in the teaching of gross anatomy in physical therapy education.* Journal of Physical Therpay Education, 1999. **13**(2): p. 41-46.
47. Nnodim, J.O., E.C. Ohanaka, and C.U. Osuji, *A follow-up comparative study of two modes of learning human anatomy: By dissection and from prosections.* Clinical Anatomy, 1996. **9**(4): p. 258-262.
48. Newell, R.L., *Follow the Royal Road: The Case for Dissection.* Clinical Anatomy, 1995. **8**: p. 124-127.
49. Johnson, J., *Importance of dissection in learning anatomy: Personal dissection versus peer teaching.* Clinical Anatomy, 2002. **15**: p. 38-44.
50. Dinsmore, C.E., S. Daugherty, and H.J. Zeitz, *Teaching and Learning Gross Anatomy: Dissection, Prosection, or "Both of the Above?".* Clinical Anatomy, 1999. **12**: p. 111-114.
51. Alexander, J., *Dissection versus prosection in the teaching of anatomy.* Journal of Medical Education, 1970. **45**: p. 600-606.
52. Nnodim, J.O., *A Controlled Trial of Peer-Teaching in Practical Gross Anatomy.* Clinical Anatomy, 1997. **10**: p. 112-117.
53. Yeager, V.L. and P.A. Young, *Peer Teaching in Gross Anatomy at St. Louis University.* Clinical Anatomy, 1992. **5**: p. 304-310.
54. Bristol Royal Infirmary Inquiry. *The Kennedy Inquiry*. 2001 [cited 17 June 2010]; Available from: http://www.bristol-inquiry.org.uk.
55. *The Redfern Inquiry*. [cited 17 June 2010]; Available from: http://www.theredferninquiry.co.uk.
56. The Stationery Office- HM Inspector of Anatomy, *Isaacs report: The investigation of events that followed the death of Cyril Mark Isaacs*. 2003, Crown Copyright.
57. Forsyth, L., B. Brittan, and M. Woof, *The implications of the Human Tissue Act 2004 for dentistry.* British Dental Journal, 2006. **201** (12): p. 790-791.
58. Royal College of Surgeons of England. *A shortage of donated bodies leave Anatomy Teaching in Crisis*. 2007 [cited; Available from: www.rcsneg.ac.uk/media/medianews/donatedbodiesanatomy.
59. Royal College of Surgeons of England. *Anatomy body donations*. 2010 [cited 10 June 2008]; Available from: www.rcsneg.ac.uk/media/medianews/donatedbodiesanatomy.
60. Patten, D., G. Finn, Editor. 2008: Durham.
61. Body Worlds. *Gunther von Hagens' Body Worlds*. 2006 [cited 17 June 2010]; Available from: www.bodyworlds.com.
62. Leibericha, P., et al., *Body Worlds exhibition—Visitor attitudes and emotions.* Annals of Anatomy, 2006. **188**: p. 567-573.
63. Preuß, D., *Body Worlds: looking back and looking ahead.* Annals of Anatomy, 2008. **190**: p. 23-32.
64. National Library of Medicine. *The Visible Human Project Factsheet*. 2004 [cited; Available from: http://www.nlm.nih.gov/pubs/factsheets/visible_human.html.
65. Acland, R., *Acland's Atlas of Human Anatomy DVD*, R. Acland, Editor. 2003, Lippincott Williams & Wilkins.

66. Whitaker, A. *Instant Anatomy*. 1999 [cited 30 June 2008]; Available from: www.instantanatomy.net.
67. Jastrow, H. and V. L, *Teaching and Learning Gross Anatomy Using Modern Electronic Media Based on the Visible Human Project.* Clinical Anatomy, 2003. **16**: p. 44-54.
68. National Library of Medicine, U.S. *The Visible Human Project*. [cited; Available from: http://www.nlm.nih.gov/research/visible/visible_human.html.
69. Donnelly, J.L., et al., *Virtual Human Dissector as a Learning Tool for Studying Cross-sectional Anatomy.* Medical Teacher, 2009. **31** (6): p. 553 – 555.
70. Robinson, A. and E.B. Jamieson, *Surface Anatomy*. 1928, Edinburgh: Oxford Medical Publications.
71. Lockhart, R.D., *Living Anatomy*. 1928, London: Faber & Faber Ltd.
72. Waterston, S.W. and I.J. Stewart, *Survey of clinicians' attitudes to the anatomical teaching and knowledge of medical students.* Clinical Anatomy, 2005. **18**(5): p. 380-384.
73. Collett, T., et al., *The Role of Living Models in the Teaching of Surface Anatomy: Some Experiences from a UK Medical School.* Medical Teacher, 2009. **31**(3): p. e90-96.
74. Barrows, H.S., P.R.Patek, and S. Abrahamson, *Introduction of the Living Human Body in Freshman Gross Anatomy.* British journal of Medical Education, 1968. **2**: p. 33-35.
75. Stillman, P.L., Jane.S.Ruggill, and D. L.Sabers, *The Use of Live Models in Gross Anatomy Teaching.* Medical Education, 1978. **12**: p. 114-116.
76. Aggarwal, R., H. Brough, and H. Ellis, *Medical Student Participation in Surface Anatomy Classes.* Clinical Anatomy, 2006. **19**: p. 627-631.
77. Rees, C.E., *The influence of gender on student willingness to engage in peer physical examination: the practical implications of feminist theory of body image.* Medical Education, 2007. **41**(8): p. 801-807.
78. Rees, C., P. Bradley, and J. McLachlan, *Exploring medical students' attitudes towards peer physical examination.* Medical Teacher, 2004. **26**: p. 86-88.
79. Rees, C. and M. Shepherd, *Students' and assessors' attitudes towards students' self-assessment of their personal and professional behaviours.* Medical Education, 2005. **39**(1): p. 30-39.
80. Rees, C., et al., *'Over my dead body?': The influence of demographics on students willingess to engage in peer physical examination.* Medical Teacher, 2005. **27**(599-605).
81. McLachlan, J., et al., *Student Attitudes to Peer Physical Examination: A Qualitative Study of Changes in Expressed Willingness to Participate.* Medical Teacher, 2010. **32**: p. e101–e105.
82. Rees, C., et al., *Medical students' attitudes towards peer physical examination: Findings from an international cross-sectional and longitudinal study.* Advances in Health Sciences Education, 2009.
83. British Broadcasting Corporation. *Surgical Gown 'marked up for ops'*. 2008 [cited 01/02/08]; Article on the Incisions Gown]. Available from: http://news.bbc.co.uk/1/hi/health/7213757.stm.
84. Bailey, P. *Feature: Flex and Ply.* 2009 4 August 2009 [cited 29/3/2010]; Available from: http://www.wellcome.ac.uk/News/2009/Features/WTX056102.htm.
85. Finn, G., *The Art of Anatomy: Innovative Methods of Teaching Anatomy*, in *ASGBI Winter Meeting*. 2009: Oxford University.

86. Hintner, M. *Bodypainting*. 2008 [cited; Available from: http://www.b-painting.com/about.htm.
87. Finn, G., *Twelve tips for running a successful body painting teaching session.* Medical Teacher, In press. **32**: p. 1-4.
88. Op Den Akker, J.W., et al., *Giving Color to a New Curriculum: Bodypaint As a Tool in Medical Education.* Clinical Anatomy, 2002. **15**: p. 356-362.
89. McMenamin, P.G., *Body Painting as a Tool in Clinical Anatomy Teaching.* Anatomical Sciences Education, 2008. **1**(July): p. 139-144.
90. Finn, G. and J. McLachlan, *A qualitative study of student responses to body painting.* Anatomical Sciences Education, 2010. **3**(1): p. 33-38.
91. Cole, S.R.G.a.T.R., *The Changing Role of Dissection in Medical Education.* MSJAMA, 2002. **287**: p. 1180-1181.
92. Finn, G., et al., *Peer estimates of low professionalism correlate with low Conscientiousness Index scores.* Medical Education 2009. **43**(10): p. 960-967.
93. McLachlan, J.C., G.M. Finn, and R.J. Macnaughton, *The Conscientiousness Index: A Novel Tool to Explore Students' Professionalism.* Academic Medicine, 2009. **84**(5): p. 559-565.
94. Finn, G., J. Garner, and M. Sawdon, *"You're judged all the time!" Students' views on professionalism: A multi-centre study.* Medical Education, In press.
95. Oxford Dictionaries, *The Oxford Dictionary of English*. 2006, OUP Oxford.
96. Van De Camp, K., et al., *How to conceptualize professionalism: a qualitative study.* Medical Teacher, 2004. **26**(8): p. 696-702.
97. Stern, D.T., ed. *Measuring Medical Professionalism*. 2006, Oxford University Press: New York.
98. North, M. *The Hippocratic Oath*. Greek Medicine 2002 [cited; Available from: http://www.nlm.nih.gov/hmd/greek/greek_oath.html.
99. The World Medical Association. *Policy: World Medical Association International Code of Medical Ethics*. [cited 02/07/2008]; Declaration of Geneva].
100. The World Medical Association, *Declaration of Geneva*, T.W.M. Association, Editor. 1948: Geneva, Switzerland.
101. Jones, D.A., *The Hippocratic Oath II: The declaration of Geneva and other modern adaptations of the classical doctors' oath.* The Catholic Medical Quarterly, 2006. **Feb 2006**(02/07/2008).
102. -□ mmah, W.r.a.-S.i.h.a. *Proceeding of the First International Conference on Islamic Medicine*. International Conference on Islamic Medicine 2004 [cited; Available from: http://openlibrary.org/b/OL2622053M/Proceeding_of_the_First_International_Conference_on_Islamic_Medicine.
103. General Medical Council. *Duties of a doctor*. 2006 [cited 10/06/2008]; Available from: http://www.gmc-uk.org/guidance/good_medical_practice/duties_of_a_doctor.asp.
104. Bligh, J., *Professionalism.* Medical Education, 2005. **39**(1): p. 4-4.
105. Peabody, F.W., *The Care of the Patient.* JAMA, 1927. **88**(12): p. 877-882.
106. Saultz, J.W.M.D., *Viewpoint: Are We Serious About Teaching Professionalism in Medicine? [Miscellaneous].* Academic Medicine, 2007. **82**(6): p. 574-577.
107. Dannefer, E.F., et al., *Peer assessment of professional competence.* Medical Education, 2005. **39**(7): p. 713-722.
108. Schuwirth, L. and C.v.d. Vleuten, *How to design a useful test* Understanding Medical Education, ed. Association for the Study of Medical Edcuation. 2006, Edinburgh: Association for the Study of Medical Edcuation.

109. Papadakis, M.A., et al., *Disciplinary Action by Medical Boards and Prior Behavior in Medical School.* N Engl J Med, 2005. **353**(25): p. 2673-2682.
110. Papadakis, M.A.M.D., et al., *Unprofessional Behavior in Medical School Is Associated with Subsequent Disciplinary Action by a State Medical Board.* Academic Medicine, 2004. **79**(3): p. 244-249.
111. Teherani A, et al., *Domains of unprofessional behavior during medical school associated with future disciplinary action by a state medical board.* Academic Medicine, 2005. **80(10 Suppl)**: p. S17-20.
112. Thistlethwaite, J. and J. Spencer, *Professionalism in Medicine*. 2008, Abbingdon: Radcliffe Publishing Ltd.
113. Arnold, L., *Assessing Professional Behaviour: Yesterday, Today and Tomorrow.* Academic Medicine, 2002. **77**(6): p. 502-515.
114. Ramsey, P.G., et al., *Use of peer ratings to evaluate physician performance.* JAMA, 1993. **269**(13): p. 1655-1660.
115. Kane, J.S. and E.E. Lawler, *Methods of peer assessment.* Psychological Bulletin, 1978. **85**(3): p. 555-586.
116. Arnold, L.P., et al., *Can There Be a Single System for Peer Assessment of Professionalism among Medical Students? A Multi-Institutional Study.* Academic Medicine, 2007. **82**(6): p. 578-586.
117. Arnold, L., et al., *Medical students' views on peer assessment of professionalism.* Journal of General Internal Medicine, 2005. **20**(9): p. 819-824.
118. Norcini, J.J., *Peer assessment of competence.* Medical Education, 2003. **37**(6): p. 539-543.
119. McCormack, W.T.P., et al., *Peer Nomination: A Tool for Identifying Medical Student Exemplars in Clinical Competence and Caring, Evaluated at Three Medical Schools.* Academic Medicine, 2007. **82**(11): p. 1033-1039.
120. McCormack, W.T., *Can peer assessment serve as a tool to provide consistent measures of humanistic qualities?* 2008, Presented at Arnold P. Gold Foundation Barriers Conference "Enhancing the Culture of Medical Education: Assessing Humanistic Growth and Mission", January 17-19, 2003 in New York, NY p. Powerpoint.
121. Arnold, L. and D. Stern, *Content and Context of Peer Assessment*, in *Measuring Medical Professionalism*, D. Stern, Editor. 2006, Oxford University Press: New York. p. 175-194.
122. Van Rosendaal, G.M. and P.A. Jennett, *Resistance to peer evaluation in an internal medicine residency.* Academic Medicine, 1992. **67**(1): p. 63.
123. Ward, P., *360-degree feedback*. 1997, London: Institute of Personnel and Development.
124. Foster, C.A. and M.R.F. Law, *How Many Perspectives Provide a Compass? Differentiating 360-Degree and Multi-Source Feedback.* International Journal of Selection and Assessment, 2006. **00014**(00003): p. 288-292.
125. Davies, H. and J. Archer, *Multi source feedback: development and practical aspects.* Clinical Teacher, 2005. **00002**(00002): p. 77-82.
126. Garbett, R., et al., *Developing a qualitative approach to 360-degree feedback to aid understanding and development of clinical expertise.* Journal of Nursing Management, 2007. **00015**(00003): p. 342-348.
127. Institue of Personnel and Development. *360 degree feedback*. 1997 [cited 21/05/2008]; Available from: http://www.cipd.co.uk/subjects/perfmangmt/appfdbck/360fdbk?cssversion=printable.

128. Lockyer, J., et al., *A study of a multi-source feedback system for international medical graduates holding defined licences.* Medical Education - Oxford, 2006. **00040**(00004): p. 340-348.
129. Stephenson, D. *360° Feedback – How Well are we Preparing Medical Students in the UK for the Real World?* [cited 27/05/2008]; Available from: www.leeds.ac.uk/medicine/meu/clinteach06/presentations/Dawn%20Stephenson.ppt
130. Stephenson, A.E., L.E. Adshead, and R.H. Higgs, *The teaching of professional attitudes within UK medical schools: reported difficulties and good practice.* Medical Education, 2006. **40**(11): p. 1072-1080.
131. Murphy, D.J., D. Bruce, and K.W. Eva, *Workplace-based assessment for general practitioners: using stakeholder perception to aid blueprinting of an assessment battery.* Medical Education, 2008. **42**(1): p. 96-103.
132. Rhoton, M.F., *A new metod to evaluate clinical performance and critical incidents in anaesthesia: quantification of daily commetns by teachers.* Medical Education, 1989. **23**: p. 280-289.
133. Rhoton, M.F., *Influence of Anesthesiology Residents' Noncognitive Skills on the Occurrence of Critical Incidents and the Residents' Overall Clinical Performances.* Academic Medicine, 1991. **66**(6): p. 359-361.
134. Rhoton, M.F., *Professionalism and Clinical Excellence among Anesthesiology Residents.* Academic Medicine, 1994. **69**(4): p. 313-315.
135. Hodges, D., J.C. McLachlan, and G.M. Finn, *Exploring reflective 'critical incident' documentation of professionalism lapses in a medical undergraduate setting.* BMC Medical Education, 2009. **9**: p. 44.
136. Papadakis, M.A. and H. Loeser, *Uisng critical incident reports and longitudinal observations*, in *Measuring medical professionalism*, D. Stern, Editor. 2006, Oxford University Press: New York. p. 159-173.
137. Dewi, S. and T. Achmad, *Optimising feedback using the mini-CEX during the final semester programme.* Medical Education, 2010. **44**: p. 509.
138. Intercollegiate Surgical Curriculum Programme, *Guidance for using the Mini-Clinical Evaluation Exercise (Mini-CEX)*. 2007.
139. Norcini, J., *The Mini Clinical Evaluation Exercise (mini-CEX).* The Clinical Teacher, 2005. **2**(1): p. 25-30.
140. Norcini, J., L. Blank, and G. Arnold, *Examiner differences in the mini-CEX.* Advances in Health Sciences Education, 1997. **2**: p. 27-33.
141. Hill, F., et al., *Implementing the undergraduate mini-CEX: a tailored approach at Southampton University.* Medical Education, 2009. **43**: p. 326-334.
142. Norcini, J., *Faculty observations of student professional behaviour*, in *Measuring Medical Professionalism*, D. Stern, Editor. 2006, Oxford University Press: New York. p. 147-158.
143. Norcini, J., et al., *The mini-CEX: a method for assessing clinical skills.* Ann Intern Med, 2003. **138**: p. 539-543.
144. Cook, D., et al., *Effect of rater training on reliability and accuracy of mini-CEX scores: A randomized control trial.* Journal of General Internal Medicine, 2008. **24**(1): p. 74-79.
145. Cook, D. and T. Beckman, *Does scale length matter? A comparison of nine- versus five-point rating scales for the mini-CEX.* Advances in Health Sciences Education, 2009. **14**: p. 655-664.
146. Kogan, J., L. Bellini, and J. Shea, *Feasability, reliability and validity of the mini-clinical evaluation exercise (mCEX) in a medicine core clerkship.* Academic Medicine, 2003. **78**(10 supplement): p. s33-s35.

147. Stern, D.T., A.Z. Frohna, and L.D. Gruppen, *The prediction of professional behaviour.* Medical Education, 2005. **39**(1): p. 75-82.
148. Fryer-Edwards, K., L. Pinsky, and L. Robins, *The Use of Portfolios to Assess Professionalism*, in *Measuring Medical Professionalism*, D. Stern, Editor. 2006, Oxford University Press: New York. p. 213-234.
149. Roberts, C., D. Newble, and A. O'Rouke, *Portfolio-based assessments in medical education: are they valid and reliable for summative purposes?* Medical Education, 2002. **36**: p. 899-900.
150. Van der Vleuten, C., *The assessment of professional competence: developments, research and practical implications.* Advances in Health Sciences Education Theory and Practice, 1996. **1**: p. 41-67.
151. Mcmullan, M., et al., *Portfolios and assessment of competence: a review of the literature.* Journal of Advanced Nursing, 2003. **3**: p. 283-294.
152. Lynch, D.C., P.M. Surdyk, and A.R. Eiser, *Assessing professionalism: a review of the literature.* Medical Teacher, 2004. **26**(4): p. 366-373.
153. Merriam, S., *Qualitative Research: A Guide to Design and Implementation*. 2009, San Francisco, CA: Jossey-Bass.
154. Vanderstoep, S. and D. Johnston, *Research methods for everyday life: blending qualitative and quantitative approaches*. 2009, San Francisco, CA: Josey-Bass.
155. Lingard, L., M. Albert, and W. Levison, *Qualitiative Research: Grounded Theory, mixed methods and action research. .* BMJ 2008. **337**(a567): p. 459-461.
156. Lingard, L. and T.J. Kennedy, *Grounded Theory*, in *Qualitative research in medical education*. 2007, Assocation for the Study of Medical Education (ASME): Edinburgh. p. 5.
157. Creswell, J.W., *Grounded Theory Research*, in *Qualitative Inquiry & Research Design: Choosing Among Five Approaches*. 2007, Sage Publications, Inc.: California. p. 62-68.
158. Glaser, B.G. and A.L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*. 2007 ed. 1967, Piscataway, U.S.A: Aldine Transaction.
159. Kuper, A., L. Lingard, and W. Levinson, *Qualitative Research: Critically appraising qualitative research.* BMJ, 2008. **337**(a1035): p. 687-689.
160. Merriam, S., *Qualitative Research: A Guide to Design and Implementation*. 2009, Jossey-Bass: San Francisco. p. 30-32.
161. Kuper A, Lingard L, and W. Levinson, *Critically appraising qualitative research.* BMJ, 2008. **337**(687-689).
162. Lincoln, Y. and E. Guba, *Naturalistic inquiry*. 1985, Thousand Oaks, CA: Sage.
163. Hall, W. and P. Callery, *Enhancing the Rigor of Grounded Theory: Incorporating Reflexivity and Relationality.* Qual Health Res, 2001. **11**: p. 257-272.
164. Noelle-Neumann, E., *The Spiral of Silence*. 1984, Chicago: University of Chicargo.
165. Donnelly, J.L., et al., *Virtual Human Dissector as a Learning Tool for Studying Cross-sectional Anatomy*, in *Association for Medical Education in Europe*. 2007: Trondheim, Norway.
166. Dawes, T., et al., *Training improves medical student performance in image interpretation.* The British Journal of Radiology, 2004. **77**: p. 775-775.
167. Patten, D., *What lies beneath: the use of three-dimensional projection in living anatomy teaching.* The Clinical Teacher, 2007. **4**: p. 10-14.
168. Dewhurst, D. and R. Ellaway, *Virtual Learning Environments*, in *A practical guide for medical teachers*, J. Dent and R. Harden, Editors. 2005, Elsevier. p. 201.
169. Harden, R., *Independent learning*, in *A practical guide for medical teachers*, D. JA and R. Harden, Editors. 2005, Elsevier: Edinburgh. p. 137.

170. Ruiz, J., et al., *Computer animations in medical education: a critical literature review.* Medical Education, 2009. **43**(9): p. 838 - 846.
171. Fry, H., S. Ketteridge, and S. Marshall, *A Handbook for Teaching & Learning in Higher Education: Enhancing Academic Practice*. 2003, New York: RoutledgeFalmer.
172. Marton F and P. Ramsden, *What does it take to improve learning?*, in *Improving Learning: New Perspectives*, P. Ramsden, Editor. 1988, Kogan Page: London, UK. p. 268–286.
173. Pandey, P. and C. Zimitat, *Medical students' learning of anatomy: memorisation, understanding and visualisation.* Medical Education, 2007. **41**(1): p. 7-14.
174. Sawdon, M.A., *Improving Knowledge Retention using KEEpad.* Medical Education, 2009(May).
175. Karpicke, J.D. and H.L. Roediger III, *The Critical Importance of Retrieval for Learning.* Science, 2008. **319**(February): p. 966-968.
176. Larsen, D.P., A.C. Butler, and H.L. Roediger III, *Test-enhanced learning in medical education.* Medical Education, 2008. **42**(10): p. 959-966.
177. Bloom, B. and G. Hastings, *Handbook on formative and summative evaluation of students' learning*. 1971, New York: McGraw-Hill.
178. Bradley, P., *The history of simulation in medical education and possible future directions.* Medical Education, 2006. **40**(3): p. 254-262.
179. Jason, H., *Claude Bernard distinguished lecture. Becoming a truly helpful teacher: Considerably more challenging, and potentially more fun, than merely doing business as usual. .* Adv Physiol Educ, 2007. **31**: p. 312–317.
180. Ogilvie, R.W., Thomas C Trusk, and A.V. Blue, *Students' attitudes towards computer testing in a basic science course* Medical Education, 1999. **33**(11): p. 828–831.
181. Finn, G., *The Colour of Your Inside*, in *Create*. 2008: University of Ulster, Belfast.
182. Finn, G., *The Colour of Your Inside: Anatomical Body Painting*, in *Experimental Biology: American Association of Anatomists*. 2009: New Orleans.
183. Finn, G.M., *The Art of Anatomy: Innovative Teaching Methods.* Journal of Anatomy, In press.
184. Eisendrath, D., *A Text-book of Clinical Anatomy*. 1 ed. 1904, Philadelphia, PA: W.B. Saunders.
185. Raftery, A.T., *Anatomy teaching in the UK.* Surgery, 2006. **25**: p. 1-2.
186. McHanwell S, et al., *A core syllabus in anatomy for medical students - Adding common sense to need to know.* Europen Journal of Anatomy, 2007. **11**(Supplement 1): p. 3-18.
187. Finn, G.M., D. Patten, and J.C. McLachlan, *The Impact of Wearing Scrubs on Contextual Learning in Undergraduate Medical Students.* Medical Teacher, 2010. **32** (5): p. 381-384
188. Verhoeven BH, et al., *Growth of medical knowledge.* Medical Education, 2002. **36**: p. 711-717.
189. Stillman, P.S., Jane.S.Ruggill, and D.L. Sabers, *The Use of Live Models in the Teaching of Gross Anatomy.* Medical Education, 1978. **12**: p. 114-116.
190. Friere, P., *Education: the practice of freedom.* 1973, London.
191. Harden, R., *Independent learning*, in *A Practical Guide for Medical Teachers*, Dent JA and H. RM, Editors. 2009, Elsevier Churchill Livingstone: London, UK. p. 168–174.
192. Bruner, J.S., *The act of discovery.* Harvard Educational Review 1961. **31**(1): p. 21-32.

193. School of Emotional Literacy. *Emotional Literacy- The Concept*. [cited 2010; Available from: http://www.schoolofemotional-literacy.com.
194. Barbour, R.S.B., *Checklists for improving rigour in qualitative research: a case of the tail wagging the dog?* BMJ, 2001. **322**: p. 1115-1117.
195. Richardson, L., *Postmodern Social Theory: Representational Practices.* Sociological Theory, 1991. **9**(2): p. 173-179.
196. Tobin, G. and C. Begley, *Methodological rigour within a qualitative framework.* Journal of Advanced Nursing 2004. **48**(4): p. 388-396.
197. Cohen, J., *Statistical power analysis for the behavioural sciences*. 2nd ed. 1988, Hillsdale, NJ: Lawrence Earlbaum Associates.
198. Metcalf, N., et al., *Peer group models in examination instruction as an integral part of medical gross anatomy.* Medical Education, 1982. **57**: p. 641-644.
199. Ganguly, P. and L. Chan, *Living anatomy in the 21st century: how far can we go?* South East Asian Journal of Medical Education, 2008. **2**(2): p. 52-57.
200. McLachlan, J., et al., *Student Attitudes to Peer Physical Examination: A Qualitative Study of Changes in Expressed Willingness to Participate.* Medical Teacher, In press.
201. Dangerfield, P., P. Bradley, And T. Gibbs, *Learning Gross Anatomy in a Clinical Skills Course.* Clinical Anatomy, 2000. **13**: p. 444-447.
202. Cox, K., *Teaching and learning clinical perception.* Medical Education, 1996. **30**: p. 90-96.
203. Ashley, E., *Medical education - beyond tomorrow? The new doctor - Asclepiad or Logiatros?*.* Medical Education, 2000. **34**: p. 455-459.
204. Godden, D. and A. Baddeley, *When does context influence recognition memory?* British Journal of Psychology, 1980. **71**(1): p. 99.
205. Godden, D.R. and A.D. Baddeley, *Context-Dependent Memory In Two Natural Environments: On Land And Underwater.* British Journal of Psychology, 1975. **66**(4): p. 325-331.
206. Rolfe, I. and R. Sanson-Fisher, *Translating learning principles into practice: a new strategy for learning clinical skills*
Medical Education, 2002. **36**: p. 345-352.
207. Richardson, H. *University budget cuts revealed*. BBC News 2010 [cited; Available from: http://news.bbc.co.uk/1/hi/education/8491729.stm.
208. Chastonay, P. Brenner, and S.G. Peel, JJ, *The need for more efficacy and relevance in medical education.* MEDICAL EDUCATION, 1996. **30**: p. 235-238.
209. Novak, J., *Applying Learning Psychology and Philosophy of Science to Biology Teaching*
The American Biology Teacher, 1981. **43**(1): p. 12-20 +42.
210. Ausubel, D. *David P. Ausubel*. [cited 21/10/10]; Available from: http://www.davidausubel.org/
http://tip.psychology.org/ausubel.html.
211. Kearsley, G. *Subsumption theory (D. Ausubel)*. Exploration in Learning and Instruction: The Theory Into Practice Database 2010 [cited 21/10/10]; Available from: http://tip.psychology.org/ausubel.html.
212. Finn, G.M., D. Patten, and J.C. McLachlan, *The Impact of Wearing Scrubs on Contextual Learning in Undergradute Medical Students.* Medical Teacher, In press.
213. Eraut, M., *Learning in the Workplace. Research Summary for House of Commons Committee.* 2000.
214. Bligh, J., *What's the use of lectures?* 5th edition ed. 1998, Exeter: Intellect.

215. Fry, H., S. Ketteridge, and S. Marshall, *A Handbook for Teaching and Learning in Higher Education*. 2000, London: Kogan Page.
216. Gordon, J.A., et al., *"Practicing" Medicine without Risk: Students' and Educators' Responses to High-fidelity Patient Simulation.* Academic Medicine, 2001. **2001**(5): p. 469-472.
217. Older, J., *Anatomy: A must for teaching the next generation.* The Surgeon, 2004. **2**: p. 79-90.
218. Lave, J. and E. Wegner, *Situated Learning: Legitimate Peripheral Participation*. 1991, New York: Cambridge University Press.
219. Daters, C., *Importance of Clothing and Self-Esteem Among Adolescents.* Clothing and Textiles Research, 1990. **8**(3): p. 45-50.
220. Kaiser, S., *Toward a Contextual Social Psychology of Clothing: A Synthesis of Symbolic Interactionist and Cognitive Theoretical Perspectives.* Clothing and Textiles Research, 1983. **2**(1): p. 1-10.
221. Feinberg, R., L. Mataro, and W. Jeffrey Burroughs, *Clothing and Social Identity.* Clothing and Textiles Research, 1992. **11**(1): p. 18-23.
222. Buckley, H. and M. Roach, *Clothing as a nonverbal communicator of social and political attitudes.* Family and Consumer Sciences Research Journal, 1974. **3**: p. 94-102.
223. Finn, G. and M. Sawdon, *Does peer and self-assessment correlate to the use of the conscientiousness index when evaluating professionalism in medical students?* , in *Association for the Study of Medical Education*. 2009: Edinburgh.
224. McLachlan, J., G. Finn, and R. Macnaughton, *The Conscientiousness Index: An objective scalar measure of conscientiousness correlates to staff expert judgements on students' professionalism*, in *Association for the Study of Medical Education (Medical Students and Professional Behaviour)*. 2008: London.
225. McLachlan, J., G. Finn, and R.J. McNaughton, *The Conscientiousness Index: an objective scalar measure of conscientiousness correlates to staff expert judgements on students' professionalism.* Academic Medicine, 2009. **84**(5): p. 559-565.
226. McLachlan, J., G. Finn, and M. Sawdon, *The relationship between conscientiousness and professionalism* in *Association for the Study of Medical Education*. 2010: Cambridge.
227. McLachlan, J., G. Finn, and M. Sawdon, *"With all due diligence": measuring conscientiousness in complex settings*, in *Ottawa*. 2010: Miami.
228. McLachlan, J., et al., *Exploring the relationship between professionalism and conscientiousness* in *Association for the Study of Medical Education in Europe*. 2009: Malaga.
229. McLachlan, J., et al., *Assessment of conscientiousness and its relation to professionalism*, in *Teaching and Learning Conference: Improving the Experience*. 2008: Newcastle.
230. General Medical Council, *Good Medical Practice*. 2006.
231. Schuwirth, L.W.T. and C.P.M. Van de Vleuten, *How to design a useful test: The principles of assessment.* ASME Series Understanding Medical Education. 2006, Edinburgh: ASME.
232. Stern, D.T., *Measuring Medical Professionalism*. 2006, New York: Oxford University Press.
233. Papadakis, M.A., et al., *Unprofessional Behavior in Medical School Is Associated with Subsequent Disciplinary Action by a State Medical Board.* Academic Medicine, 2004. **79**(3): p. 244-249.

234. Papadakis, M.A., et al., *Disciplinary Action by Medical Boards and Prior Behavior in Medical School.* New England Journal of Medicine, 2005. **353**(25): p. 2673-2682.
235. Teherani, A., et al., *Domains of unprofessional behavior during medical school associated with future disciplinary action by a state medical board.* Academic Medicine, 2005. **80**(10 suppl): p. S17-20.
236. Papadakis, M.A., et al., *Performance during internal medicine residency training and subsequent disciplinary action by state licensing boards.* Annals of Internal Medicine, 2008. **148**: p. 869-876.
237. Ainsworth, M.A. and K.M. Szauter, *Medical student professionalism: are we measuring the right behaviours? A comparison of professional lapses by students and physicians.* Academic Medicine, 2006. **81**(10): p. S83-S86.
238. *Good Medical Practice*. Duties of a Doctor. 2006: General Medical Council.
239. Arnold, L.P., et al., *Can There Be a Single System for Peer Assessment of Professionalism among Medical Students? A Multi-Institutional Study.* Academic Medicine, 2007. **82**(6): p. 578-586.
240. Arnold, L., et al., *Medical Students' Views on Peer Assessment of Professionalism.* Journal of General Internal Medicine 2005. **20**: p. 819-824.
241. Arnold, L. and D. Stern, *Content and Context of Peer Assessment*. Measuring Medical Professionalism, ed. D. Stern. 2006, New York: Oxford University Press. 175-194.
242. Dannefer, E.F., et al., *Peer assessment of professional competence.* Medical Education, 2005. **39**(7): p. 713-722.
243. McCormack, W.T. *Can peer assessment serve as a tool to provide consistent measures of humanistic qualities?* in *Arnold P. Gold Foundation Barriers Conference "Enhancing the Culture of Medical Education: Assessing Humanistic Growth and Mission"*. 2003. New York.
244. McCormack, W.T.P., et al., *Peer Nomination: A Tool for Identifying Medical Student Exemplars in Clinical Competence and Caring, Evaluated at Three Medical Schools.* Academic Medicine, 2007. **82**(11): p. 1033-1039.
245. Ramsey, P.G., et al., *Use of Peer Ratings to Evaluate Physician Performance.* Journal of the American Medical Association, 1993. **269**(13): p. 1655-1660.
246. Cohen, J., *Statistical power analysis for the behavioural sciences*. 2 ed. 1988, Hillsdale, NJ: Lawrence Earlbaum Associates.
247. Colliver, J.A., et al., *The prognostic value of documented professional behaviour in medical school records for predicting and preventing subsequent medical board disciplinary action: the Papadakis studies revisited.* Teaching and Learning in Medicine, 2007. **19**(3): p. 213-215.
248. Kruger, J. and D. Dunning, *Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments.* Journal of Personality and Social Psychology, 1999 **77**(6): p. 1121-1134.
249. Arnold, L., et al., *Medical Students' Views on Peer Assessment of Professionalism.* Journal of General Internal Medicine, 2005. **20**: p. 819-824.
250. University of Bristol. *Bristol Online Surveys*. [cited 2009 March]; Available from: http://www.survey.bris.ac.uk/.
251. Kruger, J. and D. Dunning, *Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments.* Journal of Personality and Social Psychology. 1999. **77**(6): p. 1121-1134.
252. Boud, D., *Avoiding the traps: seeking good practice in the use of self-assessment and refelction in professional courses.* Social Work Education, 1999. **18**(2): p. 121-132.

253. Eva, K.W., et al., *How Can I Know What I Don't Know? Poor Self Assessment in a Well-Defined Domain.* Advances in Health Sciences Education 2004. **9**: p. 211-224.
254. Eva, K.W. and G. Regehr, *Self-Assessment in ther Health Professions: A Reformulation and Research Agenda.* Academic Medicine, 2005. **80**(10): p. S46-S54.
255. Eva, K.W. and G. Regehr, *"I'll Never Play Professional Football" and Other Fallacies of Self-Assessment.* Journal of Continuing Education in the Health Professions, 2008. **28**(1): p. 14-19.
256. Langendyk, V., *Not knowing that they do not know: self-assessment accuracy of third-year medical students.* Medical Education, 2006. **40**: p. 173-179.
257. Davis, D., et al., *Accuracy of Physician Self-assessment compared with observed measures of competence.* JAMA, 2006. **296**(9): p. 1094-1102.
258. Finn, G. and M. Sawdon, *Students' views on peer evaluation and professionalism: "Getting away with it".* in *AMEE*. 2009: Malaga.
259. Finn, G. and M. Sawdon, *Students' views on peer assessment and professionalism: knowing when to "switch it on"*, in *ASME*. 2009: Edinburgh.
260. Finn, G., J. Garner, and M. Sawdon, *"You're judged all the time!" Students' views on professionalism: a mutlicentre study.* Medical Education, 2010. **44**: p. 814-825.
261. Van de Camp, K., et al., *How to conceptualize professionalism: a qualitative study.* Medical Teacher, 2004. **26**(8): p. 696-702.
262. General Medical Council, *Good Medical Practice* 2006.
263. General Medical Council, *The Duties of a Doctor.* 2006.
264. Reis, D.C., *Who am I and why am I here? Professionalism research through the eyes of a medical student.* Academic Medicine: Journal of the Association of American Medical Colleges, 2008. **83**(10 Suppl): p. S111-2.
265. Karniell-Miller, O., et al., *Medical students' professionalism narratives: a window on the informal and hidden curriculum.* Academic Medicine, 2010. **85**: p. 124-133.
266. Park, J., et al., *Observation, reflection and reinforcement: surgery faculty members' and residents' perceptions of how they learnt professionalism.* Academic Medicine 2010. **85**: p. 134-139.
267. Chretien, K.C., et al., *Online posting of unprofessional content by medical students.* Journal of the American Medical Association, 2009. **302**(12): p. 1309-1315.
268. Jain, S., *Practicing medicine in the age of Facebook.* The New England Journal of Medicine, 2009. **361**(7): p. 649-651.
269. Lingard, L., et al., *The rhetoric of rationalization: How students grapple with professional dilemmas.* Academic Medicine, 2001. **76**(10): p. S45-S47.
270. Ginsburg, S., G. Regehr, and L. Lingard, *To be and not to be: the paradox of the emerging professional stance.* Medical Education, 2003. **37**(4): p. 350-357.
271. Jha, V., et al., *A systematic review of studies assessing and facilitating attitudes towards professionalism in medicine.* Medical Education, 2007. **41**(8): p. 822-829.
272. Reddy, S., et al., *Third year medical students' particiaption and perceptions of unprofessional behaviours.* Academic Medicine, 2007. **82**(10): p. S35-S39.
273. Howe, A., A. Barrett, and S. Leinster, *How medical students demonstrate their professionalism when reflecting on experience.* Medical Education, 2009. **43**: p. 942-951.
274. Lindberg, O., *Undergraduate socialisation in medical education: ideals of professional physicians' practice.* Learning in Health and Social Care, 2009. **8**(4): p. 241-249.
275. Baernstein, A., et al., *Learning professionalism: Perspectives of preclinical medical students.* Academic Medicine, 2009. **84**(5): p. 574-581.

276. Finn, G., et al., *Peer estimates of low professionalism correlate with low Conscientiousness Index scores.* Medical Education, 2009. **43**(10): p. 960-967.
277. Wagner, P., et al., *Defining medical professionalism: a qualitative study.* Medical Education, 2007. **41**(3): p. 288-294.
278. Hilton, S.R. and H.B. Slotnick, *Proto-professionalism: how professionalisation occurs across the continuum of medical education.* Medical Education, 2005. **39**(1): p. 58-65.
279. Whitcomb, M.E., *Medical professionalism: Can it be taught?* Academic Medicine, 2005. **80**(10): p. 883-884.
280. Coulehan, J., *Today's professionalism: Engaging the mind but not the heart.* Academic Medicine, 2005. **80**(10): p. 892-898.
281. Huddle, T., *Teaching professionalism: Is medical morality a competency?* Academic Medicine, 2005. **80**(10): p. 885-891.
282. Gukas, I., et al., *Medical students' perceptions of peer feedback on an anatomy student-selected study module.* Medical Teacher, 2008. **30**: p. 812-814.
283. Lingard, L., M. Albert, and W. Levison, *Qualitiative research: Grounded theory, mixed methods and action research.* British Medical Journal, 2008. **337**: p. 459-461.
284. Lingard, L. and T. Kennedy, eds. *Qualitative research in medical education*. Understanding Medical Education, ed. A.f.t.S.o.M.E. (ASME). 2007. 5.
285. Creswell, J.W., *Case study. Qualitative inquiry & research design: choosing among five approaches.* Sage Publications, 2007: p. 244.
286. Kuper, A., L. Lingard, and W. Levinson, *Qualitative research: Critically appraising qualitative research.* British Medical Journal, 2008. **337**: p. 687-689.
287. Tonkiss, F., *Analysing text and speech: content and discourse analysis*, in *Researching society and culture*. 2007, Sage Publications Ltd.
288. Finlay, L., *"Outing" the Researcher: The Provenance, Process, and Practice of Reflexivity.* Qual Health Res 2002. **12**: p. 531-545.
289. Roberts, B., *Micro Social Theory.* Palgrave Macmillan, 2006. **1**.
290. Burr, V., ed. *An introduction to social constructionism*. 4 ed. 1995, Routledge: London.
291. Lave, J. and E. Wegner, *Situated Learning: Legitimate Peripheral Participation.* 1991, New York: Cambridge University Press.
292. Eraut, M., *Learning in the Workplace*, in *Research Summary for House of Commons Committee*. 2000.
293. Wear, D., et al., *Making fun of patients: medical students' perceptions and use of derogatory and cynical humour in clinical settings. .* Academic Medicine, 2006. **81**: p. 454-462.
294. Wear, D., et al., *Derogatory and cynical humour directed towards patients: views of residents and attending doctors. .* Medical Education, 2009. **43**: p. 34-41.
295. White, C., et al., *A qualitative exploration of how the conflict between the formal and informal curriculum influences student values and behaviours.* Academic Medicine, 2009. **84**: p. 597-603.
296. Turner, J., *Social identification and psychological group formation. .* European Developments in Social Psychology, 1984. **2**.
297. Wright, D., *Conformists or dynamic professionals: what's the current image that students are identifying with as a result of English Allied Health Professional Higher Education programmes?* Journal of Allied Health, 2008. **37**(4): p. e338-53.
298. Epstein, A., *Ethos and Identity*. 1978, London Tavistock.

299. Wendland, C. and C. Bandawe, *A qualitative study of medical student socialization in Malawi's College of Medicine: Preclinical Training and Identity.* Malawi Medical Journal, 2007. **19**(2): p. 68-71.
300. Goffman, E., *The presentation of self in everyday life.* 1959, Garden City, NY: Doubleday;.
301. Klossner, J., *The Role of Legitimation in the Professional Socialization of Second-Year Undergraduate Athletic Training Students.* Journal of Athletic Training, 2008. **43**(4): p. 379-385.
302. Niemi, P., *Medical students' professional identity: self-reflection during the preclinical years.* Medical Education, 1997. **31**: p. 408-415.
303. Crowe, N. and S. Bradford, *'Hanging out in Runescape': Identity, work and leisure in the virtual playground.* Children's Geography, 2006. **4**(3): p. 331-346.
304. Thompson, L.A., et al., *The intersection of online social networking with medical professionalism.* Journal of General Internal Medicine, 2008. **23**(7): p. 954-7.
305. Farnan, J., et al., *The relationship status of digital media and professionalism; it's complicated.* Academic Medicine, 2009. **84**: p. 1479-1481.
306. Heim, M., *The erotic ontology of cyberspace*. Cyberspace: first steps, ed. M. Benedikt. 1992, Cambridge: MIT Press.
307. Boyes, R., *And this is me on Facebook...helping with brain surgery*, in *The Times*. 2008.
308. Cordingley, L., et al., *Undergraduate medical students' exposure to clinical ethics: a challenge to the development of professional behaviours?* Medical Education, 2007. **41**: p. 1202-1209.
309. Bradley, P., *The history of simulation in medical education and possible future directions.* Medical Education., 2006. **40**(3): p. 254-262.
310. Finn, G., D. Patten, and J. McLachlan, *The impact of wearing scrubs on contextual learning in undergraduate medical students.* Medical Teacher, In press.
311. Engel, N., J. Dmetrichuk, and A. Shanks, *Medical Professionalism: Can it and should it be measured?* Student BMJ, 2009. **17**: p. b4455.
312. Lister, S., *Medical students "disciplined for high jinks"*, in *Times*. 2009.
313. Kitzinger, J., *Qualitative research: Introducing focus groups.* BMJ 1995. **311**: p. 299-301.
314. Shrank, W.H., V.A. Reed, and C. Jernstedt, *Fostering Professionalism in Medical Education: A Call for Improved Assessment and Meaningful Incentives.* Journal of General Internal Medicine 2004. **19**: p. 887-892.
315. Papadakis, M.A., et al., *A strategy for the detection and evaluation of unprofessional behavior in medical students.* Academic Medicine, 1999. **74**(9): p. 980-90.
316. Ainsworth, M. and K. Szauter, *Medical Student Professionalism: Are We Measuring the Right Behaviors? A Comparison of Professional Lapses by Students and Physicians .* Academic Medicine, 2006. **81**(10): p. S83-86.
317. National Patient Safety Agency, *Seven steps to patient safety. An overview guide for NHS staff*. 2004: London.
318. Friedman, B., et al., *AMEE Medical Education Guide No. 24: Portfolios as a method of student assessment.* Medical Teacher, 2001. **23**: p. 535-551.
319. Misch, D.A.M.D., *Evaluating Physicians' Professionalism and Humanism: The Case for Humanism "Connoisseurs".* Academic Medicine, 2002. **77**(6): p. 489-495.
320. McLachlan, J.C. and J. McHarg, *Ethical permission for the publication of routinely collected data.* Medical Education, 2005. **39**(9): p. 944-948.

321. Papadakis, M.A.M.D., H.M.D.M. Loeser, and K. Healy, *Early Detection and Evaluation of Professionalism Deficiencies in Medical Students: One School's Approach.* Academic Medicine, 2001. **76**(11): p. 1100-1106.
322. Ginsburg, S., N. Kachan, and L. Lingard, *Before the white coat: perceptions of professional lapses in the pre-clerkship.* Medical Education, 2005. **39**(1): p. 12-19.
323. Ginsburg, S.M.D.M., et al., *The Anatomy of the Professional Lapse: Bridging the Gap between Traditional Frameworks and Students' Perceptions.* Academic Medicine, 2002. **77**(6): p. 516-522.
324. Mayer, R., *Applying the science of learning to medical education.* Medical Education, 2010. **44**(6): p. 543-549.
325. Marton, F. and R. Saljo, *Approaches to Learning*, in *The experience of learning*, Edinburgh, Editor. 1984, Scottish Academic Press.
326. Ramsden, P., *What Does It Take To Improve Learning?* , in *Improving Learning: New Perspectives*, P. Ramsden, Editor. 1988, Kogan Page: London.
327. Biggs, J., *Student Approaches to Learning and Studying*. 1987, Hawthorn, Australia: Australian Council for Educational Research Ltd.
328. Biggs, J., *From theory to practice: a cognitive systems approach.* Higher Education Research and Development (Australia), 1993. **12**(1): p. 73-75.
329. Entwistle, N. and P. Ramsden, *Understanding Student Learning* 1983, London: Croom Helm.
330. Finn, G., *Postgraduate Certificate in Teaching and Learning: Assignment 2*, in *School of Education*. 2009, Durham University.
331. Biggs, J., *Aligning Teaching and Assessment to Curriculum Objectives*. 2003: Imaginative Curriculum Project, LTSN Generic Centre.
332. Fearne, H., *Medics riled by tests of professional behaviour*, in *The Times Higher Education*. 2009.
333. Cohen, D., *Medical students should be added to the GMC register, says Judge.* British Medical Journal, 2005. **1104**.
334. UKCAT. *The United Kingdom Clinical Aptitude Test* 2010 [cited 4th May 2010]; Available from: http://www.ukcat.ac.uk/pages/details.aspx?page=background.
335. McLachlan, J., *Personal communication*. 2010.
336. Finn, G. and M. Sawdon, *The impact of feedback and personality traits on students' ability to self and peer assess*, in *ASME Annual Scientific Meeting*. 2010: Cambridge.
337. Finn, G. and M. Sawdon. *The impact of feedback and personality traits on students' ability to self and peer assess*. in *Ottawa conference*. 2010. Miami, USA: AMEE.