Contextualising Simulation: The Use of Patient-Focused Hybrid Simulation for Clinical Skills Education

Jimmy Kyaw Tun

Department of Surgery and Cancer, Imperial College London

Thesis submission for the degree of MD(Res)

Supervisors:
Professor Roger L Kneebone
Dr Fernando Bello

Declaration of Originality

I hereby declare that I am the sole author of this thesis and that all the work within this is my own, except where it is referenced or carried out in collaboration with others who are appropriately cited.

Copyright Declaration

Jimmy Kyaw Tun BSc., MBChB., MA., MRCS., FRCR

Preface

I am a doctor in training specialising in Clinical Radiology and have previously undertaken some post-graduate medical training in other specialities including surgery and intensive care medicine. My motivation for studying simulation-based education stems originally from my interest in teaching, which I first developed during the later years of my undergraduate training and house jobs. At the time was I was particularly keen to teach those junior to me the things I wished I had been taught at medical school, most of which were in fact quite basic, but were often within the hidden curriculum.

After a few years of post-graduate medical training, I was fortunate to be offered a post as Clinical Education Fellow at Whipps Cross Hospital, London. This was a job that didn't have much structure, but gave me complete freedom for self-directed personal development. On top of the daily tasks of teaching, I also gained experience in other areas, such as exam administration, curriculum design and education management. I also completed a Masters in Clinical Education at the Institute of Education, London, during which I gain core skills in research methods as well as laid the foundation to some of the theoretical and pilot research work in this thesis (Chapter 5).

It was during this time, that I developed a keen interest in simulation-based education. The rationale and principles of simulation-based education are on the surface elegant and implicit - practice on something that represents in some way what one might do in real life and one should improve. At the same time, simulation appears to overcome some of the issues with the current methods of assessing competence particularly with workplace-based assessments, which in its current state of practice seems to be far from ideal.

Over the last decade or so, there has been an explosion in the use of simulation in its various forms with development of new techniques and approaches. Many of these approaches were innovative and have in my mind no doubt been positive in terms of educational benefit. However, I also started to observe and reflect on problems raised when incorporating simulation in healthcare education. To illustrate this, I have selected a few examples from personal experience as follows.

One example that I personally encountered as a medical student was when being taught venous cannulation on part-task simulators (arm manikin) by my clinical tutors. Despite gaining proficiency of cannulation on the simulator, I found my initial success rates on real patients were low. On reflection, the simulators used may have allowed too much room for error and as a result I was still able to cannulate the simulator successfully despite suboptimal technique. In this case, I don't believe, that the problem was entirely with the simulator itself, but the lack of both mine and my tutors' awareness of the functionalities and limitations of the simulator.

Some of the substandard practice I have witnessed arises from inappropriate use of simulation technology. For example, I have seen human patient simulators (manikins) being used for the training of doctor-patient communications skills. These simulators are highly sophisticated technologically, but do not in its current form simulate the subtleties of human emotions. In these cases, it appears that new techniques and technologies have been adopted and used simply because they exist and are exciting, but without much critical thought of the desired educational goals and outcomes. What would perhaps have been more appropriate, and is in fact already commonly used is simulated patients (professional actors) who by nature of being real human beings, look, feel and if well trained, behave similarly to real patients.

Finally, I believe that a major issue with current simulation is the poor alignment between simulation and workplace-based training. For instance, an observation I have made, is that many of my nursing colleagues are not allowed to perform cannulation due to local guidelines which state that they must complete a cannulation course and be signed off as competent. As a result some of them choose to pay to attend a simulation-based training course to develop their skills in order to be "signed off". However, upon completing the course, there is often little encouragement from their own department or ward to support them to utilise these skills in their day-to-day practice. Therefore in addition to the monetary waste, there is overall little educational benefit.

During my personal inquiry into the use of simulation-based education, I was fortunate enough by chance to meet my research degree supervisors during a conference, whom at the time were showcasing some innovative approaches to simulation. A technique that was demonstrated which is the focus of my thesis is the use of a simulated patient and an inanimate physical part-task trainer (for example a venepuncture arm), which are carefully placed in such a way, that the part-task trainer appears to be part of the patient. This technique has a number of labels in the literature, but I have chosen to term as "Patient-Focused Hybrid Simulation" for the purposes of consistency in this thesis. This technique allows for the simulation of a procedural skill in the context of a patient encounter, or conversely communicating with a patient in the context of performing a procedural skill. It is an exciting concept and one that has been adopted by many simulation practitioners since its introduction. However there remain many unanswered questions regarding it use.

My aim of undertaking this thesis was to inquire the use of Patient-Focused Hybrid Simulation in order to answers some of these questions. In addition, through the process of conducting this thesis, I also aimed to enhance my personal knowledge and understanding of the role of simulation in current healthcare education and develop my personal philosophy with respect to its use.

Of note, unconventionally and unintentionally, the time it took for me to complete this thesis was substantially longer than is usually required, largely as a result of other commitments, including fulltime clinical work and the sitting of numerous professional exams, which lead to a 2-year interruption of studies and an additional year of writing. Although the actual time spent on the thesis itself compares to that of what is usually expected, the 2 year hiatus between conducting the research and writing up for submission has allowed for a greater than normal period of personal reflection, resulting in a marked shift in my own understanding and interpretation of my work and of the current literature, to that of when I originally started. Indeed, with hindsight and given the chance, I would have liked to change my approach to conducting some of the research presented in this thesis.

In light of this, throughout this thesis, I have chosen to, where I felt appropriate present more personal reflections and honest, critical insight of my research work as well as my philosophy of the use of Patient-Focused Hybrid Simulation and simulation in general.

Acknowledgements

There are many people whom I would like to thank and acknowledge, without whom this thesis could not be completed. First and foremost, I would like to thank my wife Louise (who was actually still my fiancée at the start of my research), who has continuously supported my throughout what is widely acknowledge to all doctorate students as a challenging process.

Next I would like to thank my two supervisors, Professor Roger Kneebone and Dr Fernando Bello, whom together and individually mentored me through this process both academically and personally. Roger one the one hand was open to almost any research idea imaginable allowing for endless creativity, whilst Fernando was always able to introduce a certain level of pragmatism through his understanding of the constraints and practicalities of research. Their overall approach to supervision has allowed me the freedom to pursue a programme of research that is quite unconventional to what is usually conducted in the world of medicine and biomedical sciences, allowing me to develop a more personal philosophy in the use of simulation.

Next I would like to thank my colleagues and friends whom I had collaborated with throughout much of this thesis. Dr Stella Mavroveli, a research psychologist is unofficially (and unaware of being) my third mentor and is a master at all things related to research study design and statistics. Without her help, most of the studies in this paper would most certainly have suffered. Alejandro Granados is a computer scientist and has recently himself successfully defended his PhD, who worked tirelessly often in evenings and weekends to research and design much of the video recording technology required for capturing the participant's performance. Lucy Natalie Dormer who is both a nurse and a doctor, took part in a lot of the simulations playing various roles, essentially conducting the scenarios from within, and without whom the studies could not have been successfully completed.

I would like to also give a special thanks to Michelle Gatter, my then BSc student, who has now qualified for a number of years and is enjoying practicing medicine. Although she was my student, at times she acted more as my equal whilst working together on the Ventriloscope project. Without her organisational skills and enthusiasm, I would not have been able to collect such rich and interesting data, some of which was also successfully presented in her BSc dissertation.

I would also like to thank my friends and colleagues who helped with various aspects of the research from designing simulations and educational activities (Ruth Brown, Stuart Nuttall, Junaid Zaman, Monal Wadhera) to the relentless work of reviewing the countless hours of video performances of study participants (Stuart Nuttall, Luke Moore, Ananth Kadiyala, Graham Cole).

I would like to thank the London Deanery Simulation and Technology enhanced Learning Initiative (STeLI), and the Department of Surgery and Cancer together with their staff for providing the necessary support to allow me to complete my thesis.

Last but not least, I would like to thank the students and doctors who gave up their time to participate in the simulations.

Contributors to Studies

There following are named contributors to the various studies in this thesis who are referred to by their initials in the respective chapters.

Alejandro Granados (AG) – Chapter 5
Ananth Nag Kadiyala (AN) – Chapter 4
Graham Cole (GC) – Chapter 6
Junaid Zaman (JZ) – Chapter 6
Luke Stephen Proctor Moore (LSPM) – Chapter 6
Michelle Gatter (MG) – Chapter 6
Monal Wadhera (MW) – Chapter 6
Ruth Brown (RB) – Chapter 4 and 5
Stuart Nuttall (SN) – Chapter 4 and 5

Abstract

This thesis documents a research programme into the use of Patient-Focused Hybrid Simulation (PFHS) for clinical skills education. PFHS is an approach to simulating clinical skills that combines a simulated patient (SP) with a part-task trainer (PTT) embedding the simulation of procedural skills within a more holistic clinical context, potentially overcoming some of the shortcomings of single modality simulation. Although promising, there remains limited evidence supporting its use.

Two studies were conducted using a mixed-method approach. The first study was based on the simulation of the management of a traumatic skin laceration and consisted of two parts: 1) investigating the use of PFHS as a means of introducing clinical challenge by modifying the clinical context in which a procedure is performed; 2) exploring clinician's perception of the use of PFHS and PTT for assessing of clinical competence. These findings suggest that by changing the clinical context in which a procedure is performed, PFHS can potentially be used to objectively simulate challenge. It also demonstrated that PFHS when compared to PTT simulations was better able to induce authentic clinical behaviour within the simulation. Central to this is the presence of a human being (SP).

The second study compared the use of PFHS to patients for the training and assessment of cardiovascular examination skills. Within the limitations of this study, no significant difference was observed between PFHS and real patient-trained students in terms of their post-training performance of cardiovascular examination on real patients. There also appeared to be degree of concurrent validity between assessment of competency with PFHS and with real patients when conducted as an Objective Structured Clinical Examination (OSCE).

The work presented provides additional evidence to the existing literature to support the use of PFHS in clinical skills education. However, it also raises a multitude of questions particularly of how PFHS as well as simulation in general should be used and future directions for simulation research.

Relevant Academic Outputs

Publications:

Kyaw Tun, J., Alinier, G., Tang, J., & Kneebone, R. L. Redefining simulation fidelity for healthcare education. *Simulation & Gaming* (April 2015). 1046878115576103.

Kyaw Tun, J., Granados, A., Mavroveli, S., Nuttall, S., Kadiyala, A. N., Brown, R., Bello, F., and Kneebone, R. L. "Simulating Various Levels of Clinical Challenge in the Assessment of Clinical Procedure Competence." *Annals of Emergency Medicine* 60, no. 1 (July 2012): 112–120.e5.

Kyaw Tun, J., and Kneebone, R. L. "Bridging Worlds: Applying the Science of Motor Learning to Clinical Education." *Medical Education* 45, no. 2 (February 1, 2011): 111–114.

Kassab, E., Kyaw Tun, J., and Kneebone, R. L. "A Novel Approach to Contextualized Surgical Simulation Training." *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare* (April 10, 2012).

Kassab, E., Kyaw Tun, J., Arora, S., King, D., Ahmed, K., Miskovic, D., Cope, A., et al. "'Blowing up the Barriers' in Surgical Training." *Annals of Surgery* (July 2011): 1.

Maroothynaden, J., Tang, J. J., Granados, A., Kyaw Tun, J., Kneebone, R. L., and Bello, F. "Technology for the DS Educational Tool—facilitating Assessment, Public Engagement and Debriefing in Simulations through Technology." *British Journal of Educational Technology* 44, no. 5 (2013): E127–E132.

Book Chapter

Tang, J. J., Kyaw Tun, J., Kneebone, R. L., and Bello, F. "Distributed Simulation." In *Essential Simulation in Clinical Education*, edited by Kirsty Forrest, Judy McKimm, and Simon Edgar, 196–212. John Wiley & Sons, Ltd, 2013.

Presentations:

Kyaw Tun, J., Tang, J. J., Kneebone, R. L. Redefining Simulation Fidelity for Healthcare Education and Research. Society in Europe for Simulation Applied to Medicine Conference, Granada, Spain. June 2011

Supervised Dissertations:

Gatter, M. Validation of the Ventriloscope® for cardiac examination. BSc Surgery and Anaesthesia, Imperial College, London. 2010/2011

Table of Contents:

	Page
Chapter 1 - Introduction to Thesis	
1.1 Introduction to Chapter	1
1.2 What is Simulation-Based Education	2
1.3 Drivers for Simulation-Based Education	6
1.4 Simulation-based Assessment and Learning Theory	11
1.4.1 Learner-Centred Training	12
1.4.2 Deliberate Practice	12
1.4.3 Experiential Learning and Reflective Practice	13
1.4.5 Situated Learning and Communities of Practice	14
1.5 Simulation and Assessment Theory	16
1.6 Patient-Focused Hybrid Simulation	17
1.6.1 Context and Authenticity in Simulation of Clinical Skills	20
1.6.2 Theoretical foundations of Authenticity and Contextualised Education	23
1.6.3 Criticisms if Authenticity and Contextualised Education	25
1.6.4 PFHS and Contextualised Simulation of Clinical Skills	27
1.7 Summary to Chapter	29
Chapter 2 - Current Evidence of Patient-Focused Hybrid Simulation in Training and Assessment	
2.1 Introduction to Chapter	30
2.1.1 Assessing Quality of Research in Simulation Based Education	31
2.2 Methods	33
2.3 Results	35
2.3.1 Effectiveness of training of clinical skills with PFHS	35
2.3.2 Current evidence of psychometric properties of PFHS assessment	37
2.4 Discussion	41
2.5 Summary to Chapter	44
Chapter 3 - Methodological Approach and Philosophical Considerations	
3.1 Introduction to Chapter	46
3.2 Qualitative and Quantitative paradigms	46
3.2.1 The Quantitative Paradigm	46
3.2.2 The Qualitative Paradigm	48
3.3 Establishing a Research Approach to Simulation-Based Medical Education	49
3.4 The Mixed-Methods Approach	50
3.5 Summary to Chapter	54
Chapter 4 - Simulating Clinical Challenge with Patient-Focused Hybrid Simulation	
4.1 Introduction to Chapter	55
4.2 Methods and Materials	57
4.2.1 Study Design and Setting	57
4.2.2 Selection of Participants	57
4.2.3 Data Collection	58

4.2.4 Simulation scenario design 59
4.2.4.1 Part-Task Trainer Scenario 59
4.2.4.2 Hybrid Simulation Scenarios 59
4.2.4.3 Development of Simulated Patient with Simulated Skin Laceration 60
4.2.5 Assessment of Performance 62
4.2.6 Data Analysis 63
4.3 Results 63
4.3.1 Demographics 63
4.3.2 Performance Ratings 63
4.3.3 Comparison between Expert and Novice Ratings 64
4.3.4 Comparison between PFHS1 and PFHS2 Ratings 65
4.4 Discussion 65
4.5 Summary to Chapter 67

Chapter 5 - Exploring Participants' Perceptions of PFHS to PTT-Based Simulations for Assessment of Clinical Skills

5.1 Introduction to Chapter 69
5.2 Methods 70
5.2.1 Study Design 70
5.2.2 Participation in Simulations 70
5.2.3 Quantitative Component - Structured Questionnaires 72
5.2.3.1 Quantitative Data Analysis 72
5.2.4 Qualitative Component - Semi Structured Interviews 72
5.2.4.1 Rationale For Using the Interpretative Phenomenological Approach 72
5.2.4.2 Data collection 74
5.2.4.3 Qualitative Data analysis 75
5.2.4.4 Self-Reflexivity 76
5.3 Results 77
5.3.1 Participants' Ratings of the Simulations 77
5.3.2 Findings From Semi-structured interviews 80
5.3.2.1 Superordinate Theme 1: Suitability and Acceptability for assessment of Clinical Competence. 80
5.3.2.2 Superordinate Theme 2: Authenticity of the Simulation Experience 80
5.3.2.3 Superordinate Theme: Challenging Simulation, Simulating Challenge? 90
5.4 Discussion 91
5.5 Summary to Chapter 99

Chapter 6 - Comparing Use of PFHS to Patients for training and Assessment of Clinical Skills

6.1 Introduction to Chapter 100
6.1.1 Rationale for Investigating PFHS through Training and Assessment of Cardiovascular Examination Skills 101

6.2 Methods and Materials 102
6.2.1 Study setting 103
6.2.2 Participants: 103
6.2.3 Training Interventions 105
6.2.4 Post-Course Performance Assessment 106
6.2.5 Volunteer Patient Participants 107
6.2.6 VS/SP Simulation Design 107
6.2.7 Learning outcomes of VS/SP trained and Real Patient Trained Students 109
6.2.7.1 Post-Training Cardiovascular Examination OSCE Performance 109
6.2.7.2 Self-rated confidence and competence in performing cardiovascular examination 110
6.2.8 Post-Course Focus Groups 110
6.2.9 Data Analyses 111
6.3 Results 111
6.3.1 OSCE Performance 112
6.3.1.1 Correlation between OSCE Scores in PFHS and Real Patient stations 114
6.3.2 Self-rated confidence and competence 116
6.3.3 Knowledge and Understanding of Cardiovascular Examination Skills Pre and Post Course 118
6.3.4 Results of Focus Groups 120
6.3.4.1 Learning with the VS/SP Simulation 121
6.3.4.2 Performing on Real patients 123
6.3.4.3 Course Issues 126
6.3.5 Summary of Unstructured Observations 128
6.4 Discussion 130
6.4.1 Limitations and Considerations 136
6.5 Summary to Chapter 137

Chapter 7 General Discussions

7.1 Introduction to Chapter 139
7.2 Summary of Thesis Findings 139
7.3 PFHS, Contextualised Simulation, Training and Curriculum Considerations. 145
7.4 PFHS and Assessment Considerations 151
7.5 Proposal for Uses of PFHS and PTT 155
7.5.1 Suggestions for use of PFHS and PTT in Training 155 155
7.5.2 Suggestions for use of PFHS and PTT in Assessment 155
7.6 Personal Reflections 156
7.6.1 Introducing Inauthenticity into Authentic Assessment for Research 157
7.6.2 Assessing performance or performance for assessment? 157
7.6.3 Limitations to observing competence 159
7.6.4 Validity and Transfer of training to clinical practice - in search for the Holy Grail? 160
7.6.5 A tool is only as good as… 162
7.7 Conclusions 163

List of Tables

Table 1.1: Examples of Different Types of Simulations, Simulators and Competency Domains 4

Figure 1.2: Matrix of Guiding Principles for Authentic Education in Healthcare Education. 26

Table 2.1: Summary of Studies Investigating Training Effectiveness of PFHS 39

Table 2.2: Summary of Studies Investigating PFHS for Assessment 40

Table 4.1:Demographics of participants 63

Table 4.2: Internal Consistency Reliability and Intraclass Correlation Coefficients for the OSATS-TSC, OSATS-GRS, and DOPS across Scenarios and Raters 64

Table 4.3: Comparison of OSATS-TSC, OSATS-GRS, and DOPS scores between experts and novices across scenarios 64

Table 4.4: Comparison of in OSATS-TSC, OSATS-GRS, and DOPS scores across HS1 and HS2 for experts and novices 65

Table 5.1: Demographics of participants 77

Table 5.2: Participants' Responses to Questionnaire Items Relating to Realism, Immersion and Suitability of PTT-Based simulations and PFHS for Assessment of Clinical Competence 78

Table 5.3: Superordinate themes, Subtheme and Sub-subthemes Derived from Analysis of Semi-structured Interviews 79

Table 6.1: Cardiovascular Examinations Skills Course Outline for RP and VS Group Students 105

Table 6.2 Demographics of Study Participants 112

Table 6.3 Students' OSCE Checklist and GRS scores for the VS/SP and Real Patient Assessment Stations 112

Table 6.4 Sub-analysis of Diagnostic Performance 113

Table 6.5 Self-ratings for Confidence and Competence of VS and RP Group 117

Students in Performing Cardiovascular Examination Skills Pre and Post Course

Table 6.6 Self-ratings for Confidence and Competence of Clinical Group Students in Performing Cardiovascular Examination Skills Post-Course 117

Table 6.7 Comparison of self-ratings between RP and VS Groups (Pre and Post-course), RP and Clinical Groups (Post-course) and VS and Clinical Groups (Post-Course) 118

Table 6.8 Content Analysis of Free-Text Responses to Knowledge and Understanding of Cardiovascular Examination Questionnaire (Q1 What is the purpose of the cardiovascular examination?) 118

Table 6.9 Content Analysis of Free-Text Responses to Knowledge and Understanding of Cardiovascular Examination Questionnaire (Q2: What are the key components of the cardiovascular system?) 119

Table 6.10 Content Analysis of Free-Text Responses to Knowledge and Understanding of Cardiovascular Examination Questionnaire (Q3: What are the key skills required to perform a cardiovascular examination competently?) 120

Table 6.11: Overview of Themes from Analysis of Focus Groups 121

List of Figures

Figure 1.1: Miller's Pyramid of competence and positioning of simulation and workplace-based assessment. 16

Figure 2.1: Systematic Review Search Strategy 34

Figure 4.1 Overview of Study for Chapter 4 and 5 55

Figure 4.2: Stratification of study participants 59

Figure 4.3: Hybrid simulation of wound laceration using wearable "prosthetic" sleeve to make wound appear to be part of the patient. 61

Figure 5.1 Participant suturing wound on part-task wound suture pad 71

Figure 6.1 Overall Map of Study 103

Figure 6.2 The Ventriloscope® and its Main Components 108

Figure 6.3 Box Plots of OSCE Checklist Performance Ratings for RP, VS and Clinical Students in VS/SP and Real Patient Stations 114

Figure 6.4 Box Plots of OSCE GRS Performance Ratings for RP, VS and Clinical Students in VS/SP and Real Patient Stations 114

Figure 6.5 Scatterplots of OSCE Checklist Scores in RP and VS/SP Stations Figure 115

Figure 6.6 Scatterplots of OSCE Checklist Scores in RP and VS/SP Stations 115

Abbreviations

CoP - Communities of Practice

DOPS - Direct Observation of Procedural Skills

GRS - Global Rating Score

IQ - Intelligence Quotient

PTT - Part-Task Trainer

PFHS - Patient-focused hybrid simulation

RP - Real Patient

SBE - Simulation-Based Education

SP - Simulated Patient (Professional actor trained to portray a patient)

OSATS - Objective Structured Assessment of Technical Skills

OSCE - Objective Structured Clinical Examination

TSC - Task Specific Checklist

VS – Ventriloscope

WPBA- Workplace-Based Assessments

CHAPTER 1 - Introduction to Thesis

1.1 Introduction to Chapter

Simulation-based education (SBE) is increasingly used and adopted throughout healthcare education across a range of clinical disciplines (Mcgaghie et al., 2010; Kyaw Tun and Kneebone, 2011). The reasons for this are largely in response to a range of challenges currently faced in healthcare training and patient care. The importance of simulation in healthcare education is such that it has been advocated in a number of key polices and recommendations by government organisations and training bodies over the last decade (Department of Health, 2008; RCR, 2012).

With this rapid adoption of SBE, there has also be an explosion in the development of new approaches, techniques and technology in this field. In addition, there has been an exponential increase in the research conducted in this area, which have highlighted many areas that require further inquiry (Issenberg et al., 2011; Mcgaghie et al., 2010). One technique that has emerged over the last decade which has gained significant interest and adoption in healthcare education and which is the key focus of this thesis is "Patient-Focused Hybrid Simulation" (PFHS) (Issenberg et al., 2005). PFHS is a multimodal simulation technique that combines the simultaneous use of simulated patients (SP) with physical simulators. This technique allows participants to produce an integrated performance of both the technical aspects of a procedural skill whilst communicating and developing rapport with the patient (Kneebone et al., 2006). The theoretical and philosophical underpinnings of this technique are to provide contextualised education in recognition of the potential importance of professional context and authenticity in simulation (Issenberg et al., 2005). At the same time, it aims to place the "patient" in the centre of the simulation-based training and assessment of clinical skills (Kneebone et al., 2006).

In this chapter I will first discuss in greater detail what simulation is in the context of healthcare education and clarify some of the key terminology. I will discuss the key drivers for the increasing use of SBE particularly in response to the challenges currently faced in medical training. I will then discuss the theories that underpin the use of simulation in training and assessment. Finally, I will discuss the concept of PFHS, its theoretical foundations and rationale for this simulation technique.

1.2 What is Simulation-Based Education

There is currently a plethora of terminology that is used within the simulation literature. There also appears to be much variability and confusion regarding the use of terminology, which is unsurprising given that this is a relatively new academic field. Even the fundamental terms “simulation” and “simulator” are often used interchangeably (Kyaw Tun et al., 2015). Given the current variation, I will try to explain the nature of what simulation is and propose some definitions based on the existing literature and my personal understanding, which has developed throughout the course of composing this thesis. These will be used as a source of reference throughout the rest of this thesis for reasons of consistency.

Simulation in education encompasses a wide range of training and assessment “activities” which represent real or potentially real clinical situations (P. Dieckmann et al., 2007; Kyaw Tun et al., 2015). Simulations can vary not only in terms of the situation it aims to represent, but also the extent to which is represents a situation, i.e. it can aim to simulate components of a clinical situation (part-task) or the situation in its entirety (whole task) (Bradley, 2006). Taking cardiopulmonary resuscitation as an example, a part-task may involve just simulating cardiac compressions, whereas whole task simulations would involve all aspects of cardiopulmonary resuscitation including compression, defibrillation and administration of medication.

Simulators on the other hand are the medium through which simulations can be conducted. For example, a physical plastic patient manikin (simulator) can be used for the simulation of cardiopulmonary resuscitation skills. There are a wide range of simulators currently used in healthcare simulation, the most basic of which is perhaps a pen and paper, which can be used to simulate management scenarios. Simulators can take the form of virtual reality (VR) platforms using computer software, physical synthetic simulators, animal tissue, human beings and even the human mind for mentally simulating a process (Alinier, 2007).

Simulators can be designed to allow simulation of a component of a clinical task commonly termed as part-task trainers (PTT) or the whole of a task. I make this important distinction between simulators and simulations, as the construct of a simulation will vary depending on not only the simulator, but also how it is used. A simulator can be used to simulate different situations, for instance, a simulated patient (SP), i.e. an actor who portrays a real patient, can be used for history-taking or breaking bad news. A whole-task simulator can used for both whole-task training, but also part-task if only a component of it is utilised (Kyaw Tun et al., 2015).

Simulation can be used for training and assessment of a wide range of skills across different clinical domains. By convention, these skills have been categorised under technical and non-technical skills (Nestel et al., 2011; Yule et al., 2006). Technical skills primary encompass medical expertise, clinical technique and procedural competence, examples of which may include performing venous cannulation, wound closure and interpreting an electrocardiogram. Non-technical skills have been defined as "*the cognitive, social and personal resource skills that complement technical skills and contribute to safe and efficient task performance*"(Flin et al., 2010). These include communication, teamwork, leadership, decision-making, situational awareness and professionals. Although a distinction is made between the two categories of skills in the literature, there is recognised overlap between technical and non-technical skills (Nestel et al., 2011). In addition, it is generally considered that clinical competence requires the seamless integration of these two domains of skills (Kyaw Tun and Kneebone, 2011). A selection of the examples of types of simulation, simulators and the competency domains is provided in Table 1.1.

Different approaches and modalities of simulation can be used in combination in what is termed "hybrid simulation". One of the reasons for creating hybrid simulations is to combine the desired properties of different simulation modalities into a single entity for educational purposes (Weinger, 2010). There are various forms of hybrid simulations that have been described in the literature, an example of which is PFHS - the focus of this thesis. This technique, as described earlier combines a SP with a physical task trainer, which allows participants to practice both technical and non-technical skills simultaneously in an integrated fashion (Kneebone and Nestel, 2006). For example, one can combine a wound suture pad PTT and a SP to allow trainees to practice wound closure on a “patient” thus demonstrating technical suturing ability whilst simultaneously communicating and maintain rapport with the patient.

Table 1.1 Examples of Different Types of Simulations, Simulators and Competency Domains

Simulation	Simulator	Part/Whole-task	Competency Domain
Wound closure	Skin suture pad	Part task	Manual dexterity (Technical)
Laparoscopic Cholecystectomy	VR Simulator	Whole-task	Manual dexterity (Technical)
Patient Medication Consultation	Simulated Patient	Whole-task	Communication (Non-Technical)
Venepuncture on a patient	Hybrid of Venepuncture simulator and SP	Whole-task	Manual dexterity (Technical) Communication (Non-Technical)
Major Incident Planning Exercise	Pen and Paper	Part-task	Decision Making (Non-Technical)

Another important concept in simulation is its “fidelity”, which has implications on design, cost, feasibility and educational value, though the exact relationship between simulation fidelity and training effectiveness is not fully understood (Aggarwal et al., 2010; Alessi, 1988). Like with the terms “simulator” and “simulation”, “ the term “fidelity” also appears to be plagued with confusion in the current simulation literature (Kyaw Tun et al., 2015). This confusion is well-recognised in the literature – in the early 90’s Lane and Alluisi (1992) identified over 22 different types and definitions, and the number of definitions is likely to have increased since. Fidelity is the degree to which a simulation accurately represents reality or a referenced part of reality of interest. Yet fidelity seems to have been confused with technological sophistication of simulators in the current literature (Alinier, 2007; Maran and Glavin, 2003). The reasons for this are not clear, though it may be due to a misconception that increased realism requires more sophisticated technology. However, as illustrated by the following example, the two are not necessarily related. Consider two commonly used simulators current used for simulating laparoscopic cholecystectomies are porcine cadavers and VR simulators. The current VR simulators use the latest technology in computer imaging and haptic feedback and can allow for simulation of almost the whole surgical procedure (Dunkin et al., 2007). Despite their cost and complexity, they are still limited in realism in many aspects. The haptic devices do not quite match the real world feel of the use of laparoscopic instruments and the surgical tissue planes are not wholly accurate (Våpenstad et al., 2013). A porcine gallbladder on the other hand, being real animal tissue, may in some respects be a more approximation of the real procedure, particularly with respect to dissecting tissue planes (Kwasnicki et al., 2012). Moreover, instruments used in a porcine laparoscopic cholecystectomy simulation are the same as those used in real surgical practice. Therefore a porcine simulation can provide many of the cues that represent a real cholecystectomy despite being low tech (Kyaw Tun et al., 2015).

Another common misconception is that increased realism of the simulator leads to increased fidelity. Take for example the simulation of a patient consultation using a SP. A SP is arguably in many ways one of the highest fidelity simulators we use in current simulation practice, as by nature of being real humans, they are able to provide many of the cues that a real patient provides, particularly with respect to both verbal and non-verbal communication. However, an SP can also be instructed to behave unrealistically. For instance, they may provide their story in an artificially linear and straightforward fashion to aid a novice student to learn the process of history taking. Clearly, in this example, despite the potential high realism of the simulator, i.e. the SP, the overall resulting simulation is still of low fidelity (Kyaw Tun et al., 2015).

In summary, simulation is an activity that represents a real or potentially real situation whereas simulators are the medium through which simulations can be conducted. Simulations can be designed to represent part or the whole of a referenced situation. They can also be designed to represent a referenced situation to higher or lower levels of realism, i.e. fidelity (Kyaw Tun et al., 2015).

1.3 Drivers for Simulation-Based Education

Whilst the increasing interest and the majority of literature of SBE in healthcare has arisen in the last two decades, SBE has actually had a long history in healthcare education although the training activities were not always labelled as simulations. As far back to ancient Indian times, Sushruta Samhita, who is often described as the father of Indian surgery (practised around 600BC) had documented a range of simulations including the use of animal bladders to practice making surgical incisions, animal skin for suturing, lotus stems for ligation and dummies (manikins) for bandaging (Mackenzie, 1973). The historical use of simulation is perhaps unsurprising, as its potential educational benefits seem to be implicit. In my previous career as a surgeon in training I spent many hours practicing my suturing skills on a range of items from clothes to bananas in my spare time outside of the environment of a conventional simulation laboratory. Similarly, when I first started my radiology training, I spent many hours practicing how to use the ultrasound machine, a complex piece of equipment with a multitude of dials and buttons, on my colleagues, which can be considered to be learning through simulation.

Outside of the healthcare sector within other high-risk industries such as in the military and aviation sectors, SBE had already been widely utilised for many decades and is a key aspect of their educational culture, the reason for which again are implicit (Gaba, 2004; Ziv et al., 2003). Whilst conducting research for my thesis, I had the opportunity to gain some insight into this culture when conversing with a bomb disposal expert, who simply stated that it was absolutely necessary for him to train in a simulation before he goes out in the field, as it is a matter of life or death. In essence, it was the only way he could ensure he was maximally prepared for the real situation.

There are several reasons why SBE has gained a high level of interest as an educational technique in healthcare in recent times. First, there has been a recognised decrease in workplace-based training opportunities (Roger Kneebone et al., 2006; Scalese et al., 2008). This is in part due to a significant reduction in working hours particularly in North America and Europe where the working time of doctors has become restricted to 80 and 48 hours per week respectively (Horwitz, 2011; Ilangaratne, 2011). Whilst the working time restrictions have been put in place in order to improve work-life balance and patient safety by preventing work fatigue, it has nevertheless had a significant effect on clinical experience gained. For example, in the case of UK surgical training, a consultant surgeon would have typically completed in excess of 30000 hours of in service training prior to working time restrictions, which has since greatly reduced to 6000 hours (Chikwe et al., 2004). In addition to working time restrictions, changes to the way healthcare is provided has also diminished training opportunities. Patients are increasingly managed in the ambulatory setting, with shorter inpatient stays (Scalese et al., 2008). Many interventions are performed as day cases. The result of this is a change in the type of inpatient; whereby patients admitted tend to have a higher acuity and severity of illness, making these cases less suitable for hands-on learning by relatively novice trainees.

Finally, there has been a shift in the types of procedures performed resulting in a change in training opportunities. In surgery for example, the advent of laparoscopy surgery has resulted in a reduction in the amount of open procedures that a trainee surgeon is exposed to (Shah et al., 2005). Yet, surgeons are still required to be competent in these procedures that are still performed often for complications of laparoscopic surgery. In radiology likewise, due to advances in imaging technology, conventional (invasive) diagnostic angiograms have been increasingly replaced by other non-invasive techniques, e.g. Computer Tomography Angiogram again diminishing available training (Gould et al., 2006).

On top of the change in available training opportunities, there has also been for a long time, recognition that workplace training is insufficient for acquiring certain clinical competencies and where there is no real alternative other than simulation (Bradley, 2006). These competencies tend to be for situations that are complex and high-risk, but rare to the point that clinical exposure in itself is insufficient. In radiology for example, an extremely rare but potentially life-threatening situation is a severe contrast-induced reaction (Sarwani et al., 2012; Tubbs et al., 2009). Many radiologists may not have encountered this situation in their clinical workplace and as a result, simulation-based training has been advocated to supplement workplace training. Another example of a commonly performed simulation exercise in healthcare is major incident training. Major incidents are situations which can be natural (e.g. flooding) or man-made (railway accidents) which can have give rise to multiple fatalities and injuries and place a significant burden on the local healthcare resources (Okuda et al., 2008). Simulation may be the only realistic option for training personnel to manage these situations.

Another key driver for the adoption of simulation-based training is the increasing emphasis on patient safety. A number of serious high profile medical incidents have led to an increased public awareness and media coverage over the quality of care provided by healthcare professionals (Leape et al., 1998). For example, the Bristol Royal Infirmary Inquiry, which took place in response to an unacceptably high mortality rate amongst cardiac patients, highlighted key patient safety issues in the healthcare system (Coulter, 2002). The landmark publication by the Institute of Medicine - To Err is Human: Building a Safer Health System, by Kohn et al. (2000), brought to light many of the underlying mechanisms behind serious untoward incidences in the healthcare sector.

More recently, the Francis Report which documents the inquiry into the failures at the Mid-Staffordshire NHS Trust has highlighted a range of systematic and organisation issues relating to poor patient care and safety resulting in 270 recommendations (Francis, 2013). Many of these include recommendations for education and training particularly training which does not put patients at unnecessary risk. Recommendations promoting cultural changes in the workplace were also included. Although the report does not specifically mention the use of simulation, many authors have suggested it use to address the issues and recommendations raised, particularly as a means to improve training, patient safety and to facilitate cultural and organisational change (Hinde et al., 2016; McKenzie Smith and Turkhud, 2013; Mehdi et al., 2014; Regan et al., 2014).

As a result of these, there has been a drive to ensure clinicians are adequately trained both as individuals and within teams to be better prepared to manage a range of clinical encounters include crisis situations. There has also been a change in patient expectations in the terms how and by whom they will be cared for (Millenson, 2002). Patients now expect to be cared for by practitioners of an appropriate level of training. It is also, less acceptable to practice on a patient for the first time without prior training. There is arguably a moral imperative for the use of simulation in this respect (Ziv et al., 2003).

These issues have lead to a paradigm change in medical training from time and experienced-based to outcomes-based where the emphasis is on ensuring that clinicians acquire a pre-defined set of competencies for lifelong professional practice (Leung and Diwakar, 2002). The purpose of this is to ensure future clinicians are still competent within a more focused and shortened training system. Simulation has continually been advocated as a solution to many of these issues. From a training perspective, simulation aims to represent part or the whole of a real clinical situation whether it is performing a clinical procedure or communicating with patients and team members. It thereby provides a means for practice without the presence of a real patient. From an assessment perspective, it can provide an approximation to how clinicians might perform in the real clinical world (Gaba, 2004). In the next section, I will discuss in more detail how simulation fits into current education and assessment theory.

The potential for simulation to address the issues raised above is such that its use has been recommended in many policy documents over the last decade. Within the UK, perhaps one of the most notable publications was the 2008 Chief Medical Officer report entitled "Safer Medical Practice: Manikins, Machines and Polo Mints"[1], which provides some persuasive arguments supporting the use of simulation, not just to develop skills, but also to encourage a safety culture (Department of Health, 2008). The need for simulation is echoed in more recent reports on medical education such as in Professor David Greenway's "Shape of Training" review (Greenaway, 2013). Indeed many of specialty colleges or boards within UK and USA have since produced reports stating the need to embed simulation within their respective training curricula and educational practices (Royal College of Surgeons of England, 2014; Sachdeva et al., 2008; The Royal College of Emergency Medicine, n.d.; The Royal College Of Radiologists, 2010)}.

Although there are the arguments for use of SBE are persuasive and that there has been rapid adoption of simulation into most areas of healthcare education, there are many authors who advocated the need to ensure that its use is cost-effective. Some proponents of simulation argue that that simulation is a cost-effective as it can potentially reduced the

[1] Polo mints are commonly used in simulation training of laparoscopic skills to training manipulation skills and manual dexterity.

costs from adverse events, suboptimal care and litigation. For examples, recent studies have demonstrated that up to £1 in every £12 of the UK National Health Service (NHS) budget is spent on litigation (Ker et al., 2010). The cost of preventable adverse events in the NHS is estimated to range from £1 billion to £2.5 billion per year (Frontier Economic, 2015). The estimated cost of readmitting patients to hospital after discharge is estimated to cost the NHS £2.2 billion a year (NHS Confederation, 2011). The argument raised by these proponents is that given the evidence demonstrating that simulation may reduce errors and therefore improve patient safety, that it may potentially reduce overall costs to healthcare. A study Barsuk et al (2009) did demonstrate this – their study showed an overall significant reduction in central venous line related sepsis following insertion after introducing a comprehensive simulation-based training program in their organisation which resulted in a measurable reduction in the related costs (Cohen et al., 2010).

However, despite these arguments, a recently published comprehensive systematic review highlighted the issues of poor cost-reporting in SBE in terms of number and quality of studies despite the widespread use of simulation (Zendejas et al., 2013). Yet it is a particularly pertinent issue given the current global financial climate. On the face of it, simulation appears to be relatively expensive particularly in when compared to more traditional classroom based educational activities such as lectures, tutorials. Even for more basic simulations using part-task trainers, there are costs in terms of faculty time, faculty development, and facilities (Ker et al., 2010). Unlike tutorials and lectures, which tend to have a much lower faculty to student ratio, simulation is generally more intensive in terms of faculty requirements. Much of the commonly used simulation equipment such as integrated human patient simulators can also cost several tens of thousands of pounds each (Zendejas et al., 2013). There is therefore a need to demonstrate relative cost-effective of new simulations in relation to other educational interventions, whether it is other types of simulation or educational techniques (Cook DA, 2011). The comparison of simulation to other educational interventions, e.g. tutorials can problematic in that it can be like comparing apples and oranges. A more appropriate approach may be to better match the educational intervention to the intended learning outcomes. From the point of comparing different simulation modalities, there are some studies comparing the use of relatively low cost to high cost simulations have shown equivocal outcomes (Zendejas et al., 2013). In these cases, it is unjustifiable to use a more expensive simulation technique over the other if the both have similar proven effectiveness.

The issue of cost-benefit is however not a straightforward one. From a purely economical or utilitarian point of view, one may consider that the decision on when to use simulation should be based on maximising educational outcomes whilst minimising cost. From this perspective, studies comparing educational techniques or one simulation technique to another may seem sufficient. However, this does not take into account the underpinning any issue related to patient safety, which is the immeasurable cost of patient health and life (Cohen et al., 2010). As such, it is difficult if not impossible to draw a line on an acceptable cost of simulation if it demonstrates improved patient care. Fundamentally, it is difficult to justify morally and ethically not to use simulation if it is shown to reduce patient harm.

1.4 Simulation-based Assessment and Learning Theory

There is a range of literature throughout both the medical and wider education disciplines, which aim to provide insight into how learning takes place during simulation. A number of educational theories supporting the use of simulation have been proposed, though there is no single accepted theory. This is unsurprising given the range of simulation activities that exists (Alinier, 2007). As a representation of real clinical practice, educational theories applicable to workplace learning can also explain learning in simulation.

However simulation is often not simply a faithful representation of real clinical practice. The representation is often skewed in one way or another to enhance educational value or for practical reasons. For example, in simulation, a clinical procedure maybe simplified, much like when diagrammatic representations (as opposed to highly realistic drawings) are used in textbooks in order to aid understanding (McIlrath and Huitt, 1995). Simulated clinical encounters may also be broken down into components (part-tasks) to aid novices in learning by reducing cognitive load (Maran and Glavin, 2003). On the other hand, simulations can reflect extremities of clinical practice such as the training of crisis scenarios, which are rarely encountered (Gaba et al., 2001; Reznek et al., 2003). Often scenarios take place over a much shorter timeframe in simulation than they do in real life for practical reasons. Each of these variations may give rise to different mechanisms for learning. In addition, there are a number of educational activities used alongside with simulation, which can also enhance learning, such as expert facilitation during a simulation, and debriefing and feedback of participants' performance after a simulation. The list of theories that may explain the mechanisms of how SBE works is extensive arising from a range of paradigms from behaviourist to social constructivist, many of which have a degree of overlap. Therefore I will instead present some key selected theories and concepts, which can help provide some insight into how it fits into current healthcare education.

1.4.1 Learner-Centred Training

There is a recognised tension for doctors in training, between the need to ensure that their patients received the best care, whilst at the same time ensuring that they have the learning opportunities to become fully competent practitioners. Yet as discussed above, there has been a shift away from traditional training where much of clinicians' competencies are developed through practice on patients. Simulation as a mirror to clinical practice can allow for learner-centred training whilst eliminating some of the tensions in the real workplace of patient care. Learning can take place without the worry of time constraints in the real clinical workplace (Kneebone et al., 2004).

1.4.2 Deliberate Practice

There is a strong body of evidence across different disciplines that the development of professional expertise requires sustained deliberate practice with intention to improve (Ericsson, 2007). Improvement in performance is through identifying well-defined areas of weakness allowing focused practice, in line with principles of reflective learning, which is discussed in more detail below. In order for maintenance of expertise, practice needs to be continuous and distributed, with an element of over-learning (training beyond that required for initial proficiency). This can help develop a degree of automation of performance, along the lines of the behaviourist approach to learning (Bradley and Postlethwaite, 2003a). Examples of this include skills and drills training such as in cardiopulmonary resuscitation (CPR) training, which is time critical. At the same time, it can help develop stronger cognitive ability. Much like a master chess player who is able to think many moves ahead and respond to possible moves of their opponent, an expert clinician needs to be able to anticipate changes in a patients status and make crucial decisions accordingly.

These principles of deliberate practice can be applied to a range of skills required for clinical practice. However, current workplace-based training does not always lend itself well to this. Reduced and fragmented clinical exposure means it is often difficult to undergo sustained deliberate practice. This is particularly important in procedural and surgical skills, which have a long learning curve such as endoscopic surgery (Schlachta et al., 2001). Simulation can augment workplace training by allowing clinicians to undergo focused repetitive practice of key skills.

1.4.3 Experiential Learning and Reflective Practice

The experiential model of learning, first popularised by Kolb and subsequently adopted and modified by other educationalist such as Jarvis, describes a cyclical process of learning through acquisition of experience (Jarvis and Parker, 2006, chap. 1; Kolb and Kolb, 2012)}. Kolb's original model, describes a 4-stage process of learning which involves 1) gaining experience, 2) observation and reflection or experience resulting in 3) formation of abstract concepts and hypotheses, from which the learner can 4) test the hypothesis through application of newly formed concepts (Kolb, 1984). Experience alone is insufficient, but requires reflection which is central to development of professional expertise (Schön, 1987). Reflection can be "in-action" (thinking on ones feet) or "on action", i.e. after the event. The mechanism behind reflective learning is through an individual critically examining one's own practice to seek areas of strength and weaknesses in order to understand what it takes to perform well. Reflection can be enhanced through feedback and debriefing of performance, which is particularly important when there is a lack of personal insight. However, the busy nature of clinical practice often makes this difficult. Through the process of simulating a clinical encounter and subsequent debrief, clinicians can have the opportunity first gain experience within an environment which is a proxy to the real clinical world, allowing them to then reflect and identify areas in need of improvement (Fanning and Gaba, 2007).

1.4.4 Situated learning and Communities of Practice

Situated learning is a theory first popularised by Lave and Wenger (1991), which emphasises on the importance of social practice on learning. Learning is through socialisation, visualisation and imitation. The mechanism, by which learning takes place, is through learning within a community of practice (CoP). A CoP can be considered to be group of practitioners who share common professional interests, learning and developmental requirements. For a community of practice to exist, it requires three vital components as follows (Wenger, 2011):

- *Domain*: A CoP is not only just a group of colleagues, but on that has a shared domain of interest.
- *Community*: For a CoP to form, members of the community cannot just exist, but need to engage with one another such that relationships where learning and development occur can be built.
- *Practice*: In addition to shared interests and engagement, a CoP is formed when members or practitioners collectively through practice, experience and reflection develop and learn. Strategies and solutions to problems within the domain of interest can be cultivated.

CoPs are dynamic and constantly evolving. Communities may change due to needs or evolution of the domain. Members may also come and go. Lave and Wenger (1991) describe newcomers to a CoP as legitimate peripheral participants. These new members as the names suggests, initially participate in more peripheral activities, but through experience, development of self-identity and acceptance by the CoP gradually play more central roles. Through this migration, novices gain a sense of their professional identity, which encompasses its duties, boundaries, values and aspirations (Lingard et al., 2002).

Learning and becoming a more central participant in this respect requires not just acquisition of abstract knowledge and skills but how they apply within a specific social context and culture, which in the case of healthcare is within is their designated clinical workplace (Wenger, 2000). For instance competent performance of cardiopulmonary resuscitation requires not only the understanding of the principles of human physiology and life support skills, but also how these apply in a complex and often messy environment.

Competent practice requires being able to “legitimately” contribute and participate in a team.

Simulation has the potential to contribute to situated learning in a number of ways. First, by nature of representing real clinical practice, simulation can help learners to understand and apply their knowledge and skills in the clinical context. It can facilitated the translation of de-contextualised and abstract knowledge and skills to real practice (Bradley and Postlethwaite, 2003a).

Simulation may be able to act as a platform for members-to-be to make steps towards becoming more legitimate, by providing learners with an environment that allows them to understand their role and foster professional development (Kneebone et al., 2005). This has potential benefits in for preparation to work in the modern healthcare environment which is not only complex, but due to its highly professionalised nature can seem inaccessible to the novices, a phenomenon that is well documented amongst student doctors and nurses (Levett-Jones and Lathlean, 2008; O’Brien et al., 2007).

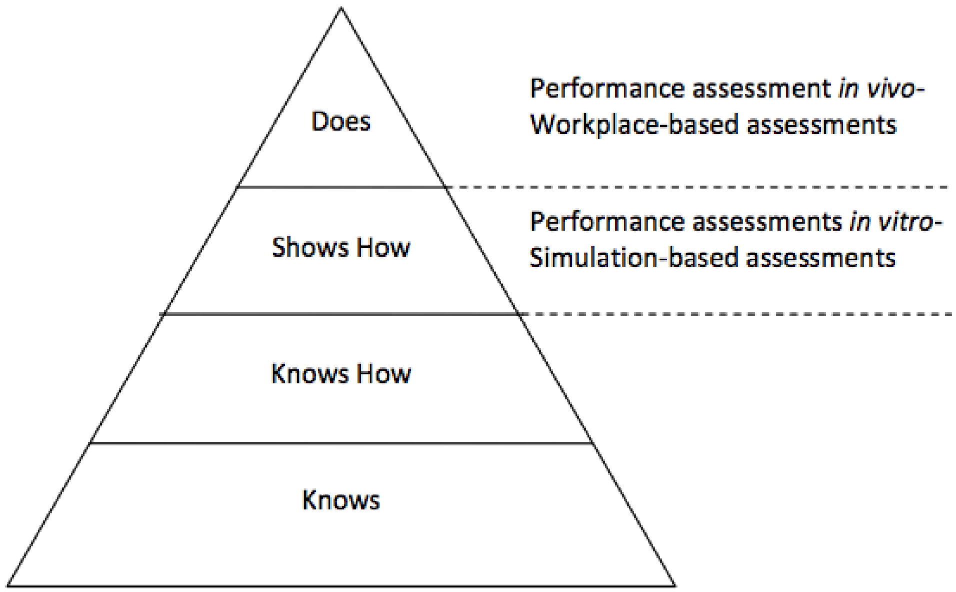
Simulation may also provide a means to foster a CoP by providing the necessary components (domain, community, practice) for situated learning in a relatively patient risk-free in-vitro environment, i.e. outside the clinical workplace. Furthermore in the healthcare environment, members are often required to rapidly integrate into a functioning team (Bezemer et al., 2016). Taking CPR again for instance, in this emergency situation, members of the resuscitation team may have not met other members of the team prior to congregating around their arrested patient. Yet they are required to rapidly become legitimate participants to function adequately together to manage the situation at hand. Simulation can provide a platform for members who may not have previously worked together to foster their shared domain and understand the requirements of practice within the professional context and culture.

In summary, in this section I described a range of theories from a range of educational paradigms that may explain how learning occurs through simulation, as well has how simulation activities can be constructed to support these approaches to learning. The importance of context which is the focus of this thesis resonates within some of these theories, particular in Lave and Wenger's situated learning.

1.5 Simulation and Assessment Theory

In addition to training, simulation is increasingly used as a form of assessment of clinical competence. There are strong reasons for using simulation for assessment. This is especially pertinent in the current outcomes-based paradigm of medical training where there is a strong focus on an individual's clinical competence. There are several frameworks described in the current literature for assessment of competence. One of the most widely used in healthcare education is described by Miller who proposes four tiers of assessment 1) knows how, 2) knows, 3) shows how and 4) does (Figure 1.1) (Rethans et al., 2002). According to this framework, the most valid method of assessment is through direct observation of real clinical performance in the workplace, i.e. workplace-based assessments (WBPA) (Norcini and Burch, 2007). WBPA are valid in the sense that they measure actual clinical performance and have been increasingly used in recent years. However, there are a number of key issues with respect to the use of WPBA (R Kneebone et al., 2006).

Figure 1.1: Miller's Pyramid of competence and positioning of simulation and workplace-based assessment. Adapted from Schuwirth and van de Vleuten (2003)



First, it relies on real clinical encounters taking place. However, clinical encounters are random by nature - one cannot predict when the next case of appendicitis may be admitted to hospital. Sampling a clinician's performance over a range of clinical scenarios may also be difficult, particularly for rare or serious cases. Secondly, the busy nature of the clinical workplace where patient care is the priority does not always permit WBPA to take place (Quantrill and Kyaw Tun, 2012).

Simulation-based assessment provides an approximation of what happens in the real clinical world. In this respect it is more authentic reflection of real clinical practice than knowledge-based assessments such as written examinations and therefore has strong face validity. According to Miller's pyramid of assessment of competence, simulation allows for assessment at the second highest tier, "shows how", by allowing clinicians to demonstrate skills "in vitro", i.e. in a controlled setting (Rethans et al., 2002).

Simulation-based assessment also allows clinicians to demonstrate competence without harm to patients, which is particularly crucial in rare and high-risk scenarios (Ziv et al., 2003). Often in medicine like with other high-risk industries, there is no alternative to training and assessment other than through the simulation of these cases. Simulations can also be standardised and therefore offers a degree of reliability (Scalese et al., 2008).

Finally, beyond the purpose of assessment as means to measure competence, it is widely acknowledged that assessment drives learning (van der Vleuten and Schuwirth, 2005). The content and structure of a given assessment affects what and how a student learns. Simulation-based assessment, by nature represents real world practice, and should therefore in theory promote the learning of skills for professional practice.

1.6 Patient-Focused Hybrid Simulation

Having discussed what simulation is and the theoretical foundations for its use, I will now focus on PFHS in greater detail and provide a background for it's theoretical and philosophical underpinnings. As mentioned earlier PFHS is a relatively new technique in simulation which combines an SP and PTT together. There are various examples that are currently in use or described in the literature, which I have summarised in Table 1.1. PFHS, as with many new techniques, has several terms to describe the same approach. Within the current literature PFHS has been termed “hybrid simulation”, “patient-focused simulation” and “scenario-based simulation” (Issenberg et al., 2005; Roger Kneebone et al., 2006). Likewise, the term “hybrid simulation” has been used to describe a range of other types of simulation activities particularly in other academic fields such as in gaming and modelling (Shanthikumar and Sargent, 1983). For reasons of consistency, I will use the term PFHS throughout the thesis.

To understanding the how PFHS fits into the simulation of clinical skills, I first propose to position it in relation to what is generally considered to be the conventional approach to simulation of clinical skill, i.e. using part-task trainers (PTT) in isolation, in terms of the degree to which these types of simulations are embedded in the clinical and professional context. PTT when used in isolation for simulation, such as a physical arm manikin used for simulating venepuncture can be considered as relatively “decontextualised” in that the task of performing venepuncture is out of the clinical context, i.e. the conditions under which the procedure is usually performed. PFHS on the other hand, by means of combining a SP with a PTT allows relatively “contextualised” and more authentic simulation. Using the simulation of venepuncture as an example again to illustrate this, a PFHS using a SP and venepuncture manikin within a clinical scenario allows participants to perform a simulated venepuncture in terms of physical (arm attached to patient) and social context (situating the procedure on a patient and under conditions where there is a clinical purpose). I described these two types of simulations in terms of being “relatively” decontextualised and contextualised as PTTs used in isolation still provide some context, but not to the same extent as PFHS does. Essentially PTT still involves placing a needle into a simulated patient’s arm and not say for instance into an orange and therefore the act of placing a needle is still embedded in some clinical context.

PFHS which aims to contextualise simulation was introduced in response some of the concerns raised regarding the state SBE of clinical skills which as mentioned earlier, conventionally focused on isolated technical skills training and assessment using PTT

(Ellaway et al., 2009). The reasons for the widespread use of part-task trainers is not certain, though it is probably related to how SBE of clinical skills has historically developed and evolved over time. From the perspective of training, as mentioned earlier, it has been acknowledged for some time that workplace-based training alone does not provide sufficient or consistent training of clinical skills (Remmen et al., 1999). As a result, a multitude of PTT were developed and clinical skills laboratories where technical skills training of these clinical procedures predominately take place were widely adopted and is now a key part of almost all modern medical schools (Bradley and Postlethwaite, 2003b; Dent, 2001). The expansion of the availability of PTTs is likely also in part driven by the simulation industry - one only needs to walk past the commercial stands in a typical medical education scientific conference to see how many different PTT products for clinical skills simulation there are currently on the market.

As well as the influence of training needs, there was a recognised need to standardise performance-based assessment of clinical skills. In the late 70's Ronald Harden who pioneered the Objective Structured Clinical Examination (OSCE), raised issues of performance-based assessments of medical student competency used at the time which was primarily in the form of long and short cases (Harden et al., 1975). These typically required students to take a history and examine a few selected patients followed by a relatively unstructured viva voce type assessment of the cases (Wass and Van Der Vleuten, 2004). Harden (1988) argued that such an approach lacked standardisation as well as adequate sampling of clinical competencies and hence proposed the use of OSCEs. OSCEs have evolved over the years and there have been several different adaptations, though the fundamental principles are similar in that they consists of a series of short (5 to 15 minute) standardised assessment stations, each of which aims to focus on limited set of competencies (Boursicot and Roberts, 2005). PTT are used extensively particularly to allow standardised assessment of competence of invasive or intimate procedures, such as digital rectal examinations in these OSCEs (Scalese et al., 2008).

From the perspective of assessment, decontextualised PTT-based simulations allow reproducible, reliable focused assessment of a well-defined competency. From the perspective of training, there are clear benefits to breaking down clinical encounters into isolated tasks to aid the learning process - expertise is developed through sustained deliberate practice as discussed earlier (Anders Ericsson, 2008). Professionals are often able to identify their own weaknesses and through the focused practice of component tasks, can improve their performance. However to improve performance in the professional context, the component tasks that were practiced and developed in isolation need to be integrated into ones practice. As Ericsson (2003, chap. 3) argues:

> *"Deliberate practice is… designed to improve specific aspects of performance in a manner that assures that attained changes can be successfully integrated into representative performance. Hence, practice aimed at improving integrated performance cannot be performed mindlessly or independent of the representative context for the target performance..."* (p. 79).

This relatively decontextualised focus on technical skills training and assessment on PTTs has no doubt played an important role in the education of clinicians and medical students over the years. However, a number of concerns have also been raised regarding this approach to clinical skills training which has been described as "reductionist" and does not necessarily take into account this issue of integration into professional context (Kneebone et al., 2006). There has also been growing recognition of the importance of context and authenticity in education within healthcare and the wider education community and is a key area that has been highlighted in the current simulation literature where future research efforts are recommended (McGaghie 2010). There are a number of theoretical reasons supporting the use of "contextualised" simulation of clinical skills with PFHS, which I will discuss below.

1.6.1 Context and Authenticity in Simulation of Clinical Skills

Before I begin my discussion of how PFHS may be educationally beneficial by means of introducing context in simulation training, I will first present the philosophical origins and theoretical foundations for authenticity and contextualised education and how this relates to PFHS.

Authentic education is the learning and assessment of knowledge and skills in contexts that reflect the way it will be useful in real life (Herrington et al., 2014). Context has been defined as, "*the circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood*"(OED). The term "authenticity" in the context of learning and assessment gained widespread adoption in the late 1980's and early 1990's and was first popularised by the likes of Archbald and Newmann (1988), and Wiggins (1990). However, the underlying principles of authentic assessment and learning have been described much earlier in the education literature by educationalists such as Fitzpatrick, Morrison and Lindquist (in Mehrens, 1992). For example, from the perspective of assessment, Lindquist (1951) argued that "*it should always be the fundamental goal of the achievement test constructer to make the elements of his test series as nearly equivalent to or as much like the elements of the criterion* (such as professional clinical practice) *series as the consequences of efficiency, comparability, economy and expediency will permit*" (p.152).

The authentic education paradigm was introduced in response to the criticisms to traditional education, which emphasises on the learning of abstract principles and standardisation of tests for fairness and reliability (Splitter, 2008). Crucially, there was increasingly recognition amongst educationalists of the shortcomings of traditional education paradigm in terms of utility, whereby many students were unable to perform in the real world despite years of academic schooling. For example, traditional assessments such as multiple-choice questions focus on the 'knowing' of facts, which are often abstract as opposed to the application of knowledge (Montgomery, 2002). The format of these tests also emphasise on structure and forced choices. Yet tasks in professional practice are often ill defined, with multiple solutions and require the application of prior knowledge.

Wiggins (1993) a pioneer of the authentic education paradigm presents a compelling case for the need for authentic education through reflections of his own teaching experiences. In his paper titled, " Assessment: Authenticity, Context and Validity" he starts his discourse of authentic assessment by putting forward his view on what performance is. He describes performance as more than just skills and drills but the integration of a range of knowledge and skills and professional judgement. Drawing from his experience a soccer coach he noted that some of his best players during practice, who were excellent at drills and set pieces (which can be considered a form of part-task training) performed poorly in a real match (Wiggins, 1993). What was missing was the ability to integrate and apply the skills learnt in practice into the context of a real football match. In other words, learning to kick a ball into the back of a net is very different from scoring a vital match-deciding penalty kick in front of a large audience (Kyaw Tun and Kneebone, 2011).

Although there has been much published on authenticity since being advocated over two decades ago, and its importance as an approach to education has been recognised, it has been noted that there is evidence to be some a degree in diversity and confusion regarding the concept of authenticity (Kreber and Klampfleitner, 2012). This is perhaps unsurprising given the adoption of authenticity into educational sectors across a range of settings.

However, in its infancy, the authentic education movement was primarily focused in the compulsory education sector. Earlier work also such as in Wiggins landmark paper mentioned above, were originally focused on making the argument for authenticity in assessment (Wiggins, 1993). Yet current literature in authentic education reflects a range of ideas across different contexts, from compulsory to post-graduate and adult education, and across multiple disciplines (Archbald and Newmann, 1988; Cranton, 2001; Darling-Hammond and Snyder, 2000; Herrington et al., 2014). Authenticity is now also generally considered in relation not only to assessment, but also instruction and learning. This is also perhaps predictable as assessment and learning are inextricably linked to one another. Indeed Wiggins in an earlier paper, argued not just for authentic assessment, but for teaching to the (authentic) test (Wiggins, 1989). Likewise Newmann's work produced around the same time did not differentiate between learning or assessment when arguing for the need for authenticity (Newmann and Wehlage, 1993). Authenticity in assessment aims to drive authenticity in learning. Likewise valid assessment of authentic learning outcomes can only achieved if the assessment themselves are authentic (Mueller, 2008) .

Some authors take authenticity in its most literal sense, i.e. "genuine" and aligned to real world problems, inferring the need for elements such as the physical (environmental) and social contexts of various real world tasks (Lombardi, 2007). From this point of view, the authenticity of an educational activity how realistic may be considered synonymous.

Simulation by representing real world activities can also be seen as an authentic educational activity. Others have emphasised on the value of authentic education to be not just the resemblance of a task in the classroom to that in the workplace in order to help develop specific competencies, but to foster individuals who can make sense of, and construct their own understanding of the world such that they become the independent professionals capable of reasoning, in-depth inquiry, and lifelong learning (Splitter, 2008). Recent work by Kreber et al (2012) go further to explore the meaning of authenticity in teaching, i.e. from an educator's perspective which encompasses sincerity, values, having a sense of care for both their subject and students, pursuing moral questions, and striving towards greater self-knowledge.

Given this diversity of concepts, to understand how the concept of authenticity and context and may apply to healthcare education and more specifically to simulation, an understanding of its theoretical foundations is required.

1.6.2 Theoretical foundations of Authenticity and Contextualised Education

The authentic education paradigm is rooted in the social and constructivist theories of education. The constructivist paradigm was first described and established by Piaget which posits that humans construct knowledge through their perception and understanding of the world (Kaufman and Mann, 2000). Learning and acquisition of knowledge occurs through interaction between new experiences and pre-existing ideas. In this respect, every learner develops his or her own unique representations of reality. Learning is therefore dependent on a complex relation between cognitive, affective and socio-cultural factors. This constructivist paradigm has given rise to Lave and Wenger's situated learning theory discussed earlier in section 1.4, and contextual learning theory, which are also strongly linked to the authentic education paradigm (Jonassen et al., 1995). Contextual learning theory suggests that learning occurs when learners process new information or knowledge in a way that makes sense to them in their own frames of reference (Hull, 1993). Learning is therefore context dependant.

Within the domain of educational psychology, the importance of context in learning and assessment of ability is increasingly recognised, though it has been suggested for sometime.

One of more widely cited classic experiments is by Goddena and Baddeley (1975) who demonstrated that divers who learned word lists underwater were better able to recall them underwater and less so on dry land; likewise, divers who learned word lists on dry land were better able to recall them on dry land and less so underwater. In essence, memory recall was better when conducted in the physical context in which something was learnt. Whilst relatively crude and limited, this experiment was nevertheless important and for the first time highlighted empirically the possible importance of context.

There are numerous examples in the psychology literature that have since raised the issue of the importance of context in learning and assessment. Brazilian street children for example have been shown to demonstrate highly levels of mathematical reasoning and ability when conducting calculations for their street business, yet when faced with a traditional classroom or school-based written test, are unable to perform. Similarly, it has been demonstrated that many students face difficulty when translating relatively decontextualised knowledge learning in the classroom to the workplace (Ceci and Roazzi, 1994).

This body of work provides some compelling arguments for the need for contextualised education. However, although closely related, arguments for contextualised learning alone are insufficient for authenticity. Whilst contextualised learning as seen from the examples above may have scope to improve the training of the task at hand, it does not necessarily achieve in full what authentic education does. The proponents of the authentic education paradigm including Wiggins and Newmann, argue that the ultimate goal of authentic education is to foster competent independent professionals (Newmann and Wehlage, 1993; Wiggins, 1993). In essence, it aims to encourage development of individuals who are reflective, self-directed and creative, such that they can cope with the multitude of nuances and subtle variations of problems presented in the real world outside the classroom.

From the perspective of assessment, there are also strong arguments supporting the need for authenticity. Gulikers et al (2004) for example, argues that authentic assessment by nature have strong validity, i.e. the assessment measures what it is supposed to measure as it is designed in relation to professional practice. Secondly, she also argues that authentic assessment has the potential to positively influence learning. By nature of reflecting real professional practice, it has the potential to promote development of competencies required for real-life professional practice. There is also evidence that as it is directly relevant to real-life practice, learner's motivation is increased. This is inline with the widely accepted notion that assessment drives learning. Messick (1994) also raises the issue of whereby validity of an assessment may be threatened by construct underrepresentation (in relation to professional practice). This issue may be resolved by increasing the authenticity of a given assessment, and providing stronger links the professional context.

As discussed earlier, the concept of authenticity is a complex one. In one respect it is aligned with the principles of social and contextualised education, though it also represents more in terms of the values it places on what it aims to achieve. Therefore rather than providing singular definitions for authentic assessment and learning, it is perhaps more appropriate to view it in terms of guiding principles. Newman et al (2015, chap. 1)} provided a useful matrix to map the authenticity of educational activities based on these principles. Although first described in relation to compulsory education, the principles appear to also be applicable to healthcare education. An adapted version of the matrix by Newmann et al in relation to healthcare education is presented in Table 1.2, which provides some of the bases for the discussions in the thesis.

1.6.3 Criticisms of Authenticity and Contextualised Education

There are a number of criticisms towards the authentic education movement and contextualised learning. Anderson et al (1996) argue against the overemphasis of contextualised learning. He points out that there is evidence that some of what is learnt in decontextualised and abstract activities are transferable to different real life contexts. The degree of transfer is skill-dependant and the context in which it is applied to. For example, he argues that skills such as reading and simple arithmetic is generally transferable across different contexts. He also argues that learning is often more effective when broken into component parts. Learners often focus on components to reduce cognitive load. In this respect, inclusion of the social and physical context of a given task may be detrimental to learning.

Figure 1.2: Matrix of Guiding Principles for Authentic Education in Healthcare Education.

	Standards for Authenticity	
	Learning	***Assessment***
Construction of Knowledge, Skills and Understanding	Promoting higher order thinking such that student are able to synthesise information to arrive at conclusions that produce new meaning and understanding for them Develop not only ability to perform clinical skills, but and understanding of different approaches and why.	Demonstrating ablity to organisation of clinical information and consideration of alternatives solutions to clinical problems Demonstrate not just competency, but critical understanding for application of clinical skills.
Disciplined Inquiry	Fostering deep knowledge and understanding of complex clinical problems Promote engagement with other clinical professional and foster habit of critical appraisal skills of ones and others' work in a way that builds an improved and shared understanding of ideas or topics	Demonstrate understanding at a conceptual level and engage in conversation with other clinical professions as well as critically appraise theirs and ones work in a way that builds an improved and shared understanding of ideas or topics.
Value Beyond "School"	Connections to the clinical workplace beyond the "classroom"* Social support within the clinical workplace for student achievement beyond the classroom.	Problems connected to the clinical workplace beyond the "classroom"* Aligned to an "audience" beyond the classroom, i.e. have relevance to real professional practice.
* "Classroom" refers to formal places of study outside the clinical workplace, e.g. clinical skills laboratory etc.		

A further argument is raised by Terwilliger (1997) who argues that using strong face validity (i.e. that it appears to be valid) as an argument for authentic education is insufficient to support the argument that it results in stronger construct validity. Indeed, despite over two decades of the existence of the authentic assessment movement, there appears to be a lack of strong empirical psychometric data to support its usage. A study by Gulikers et al (2005) comparing the use of relatively authentic to in-authentic learning environments for consultancy training demonstrated no additional value in the authentic learning environment in terms of learning outcomes measured as well as student perception and satisfaction. The authors suggested some possible reasons for their findings including unnecessary distractions arising from the authentic learning environment leading to inefficient learning and discontent amongst students (Gulikers et al., 2005).

However on further analysis, there may be other possible explanations. First, it has been suggested that conventional methods of determining educational validity may not be appropriate for authentic activities (Hodges, 2003). A closer look at the study by Gulikers et al (2005) demonstrates that the methods used to assess learning outcomes (written report and multiple choice questions), may be insufficient to measure the effect of increasing authenticity. In addition, whilst the study made attempts to provide an authentic environment by creating a more realistic simulation, it may have been lacking in other domains of authenticity such as the trying to ensure value beyond the classroom.

Others such as Terwilliger (1997) have also dismissed the claims that authentic activities generally promote higher order thinking. In particular, he raises concerns that the role of “knowledge” in assessment is downgraded, a view that is also supported by others (Hodges, 2006). He argues that there is strong evidence supporting the role of knowledge in professional practice and that there should be a multimodal strategy to education.

In this thesis, my aim is not to criticise or disprove the value of decontextualised simulations. As discussed earlier, there are strong reasons for practicing isolated clinical tasks out of contexts. Instead, the thesis aims to inquire into the use of contextualised simulation with PFHS and its potential value.

1.6.4 PFHS and Contextualised Simulation of Clinical Skills

Having presented some underlying principles of authenticity and contextualised education, I will now return to discuss how this relates to simulation based education of clinical skills. Like with Wiggin’s observation of the lack of transferability between the skills and drills exercises during practice to a real football match, from my personal experience as a learner and teacher, what is learnt in a simulation laboratory does not always directly transfer to the real clinical practice. This issue has also been raised by a number of other authors (Kneebone et al., 2004; Laschinger et al., 2008) and a comprehensive systematic review has also demonstrated that there remains a limited of evidence of transferability from the clinical skills laboratory to the real clinical workplace (Lynagh et al., 2007).

A key reason for this gap between simulation and the clinical world may be that real clinical practice requires more than just skills and drills of technical skills but the integration of these skills with other aspects of clinical competence such as communication, decision-making and professionalism (Kyaw Tun and Kneebone, 2011). When examining a patient for example, it is not only the technical expertise of eliciting clinical signs that is required, but also communication and professionalism to ensure patient compliance particularly when conducting manoeuvres which require patients to follow instructions, whilst maintaining rapport and trust (Nestel et al., 2009). This is a particularly crucial in certain situations such as when performing an intimate examination. However current decontextualised simulation of clinical skills training with PTT focuses only on technical skills and therefore may not prepare students for real clinical practice. PFHS on the other hand, by nature of embedding procedural skill simulation within a clinical scenario on a (simulated) patient, can allow students to be trained and assessed in an integrated manner.

In addition, to the issue of integrated performance, the sole practice of decontextualised component clinical tasks does not provide the full picture of how these tasks are applied in the real life context. Clark et al (1997) argue the rehearsal of skills in the laboratory removes the uniqueness of the complexity and constraints of the clinical setting. The context in which a task is performed as demonstrated in some of the psychology experiments discussed above, may however influence performance. From the perspective of clinical performance, even a seemingly simple clinical task can become challenging depending on the clinical situation - placing a peripheral cannula in a fit young compliant patient is very different from perform the same task on a shocked patient undergoing cardiopulmonary resuscitation. Procedural training should not only take into account the nominal challenge of the procedure itself, but the functional challenge. Previously, Kneebone et al (2007) proposed the need to map simulation activities according to relative complexity and clinical risk in relation to the learner. PFHS by means of allowing the modification of challenge by changing the clinical context, in which a clinical skill is performed, may allow for more tailored simulation whereby complexity can be introduced according to the learners' or assessees' educational needs (Kneebone et al., 2007). For example, a PFHS scenario for wound closure can be designed to be more complex, by designing the scenario and patient's role to portray a drunken uncooperative patient. A PFHS scenario for respiratory examination can by made more complex, by having the SP portray that they are in respiratory distress.

Lastly, the use of PFHS represents a philosophical stance that transcends educational utility in terms of the training and assessment of competence. As stated above, much of current SBE takes places in the relatively sterile environment of the clinical skills laboratory. There also appears to be an artificial divide between clinical skills training in the laboratory and the clinical workplace. Crucially, the presence of a real human being / patient is missing (Kneebone et al., 2006). A key goal of PFHS is to refocus current clinical skills simulation by means of placing a real human being in the centre, to encourage educational goals to be focused around the patient and their care.

I.7 Summary to Chapter

In this chapter, I discussed the nature of SBE and the reasons for its increasing use throughout healthcare education, which stem primarily from the changes in the culture of medical training, and which ultimately aims to ensure clinician competency and promote better patient care. Simulation-based education encompasses a range of activities that represent real or potentially real situations. Within healthcare education, these can be used as a proxy to real clinical practice to allow training and assessment of healthcare professionals and students. There are a multitude of educational theories that underpin the use of simulation in healthcare education. Along with the rapid adoption of simulation came an explosion of new approaches and techniques to SBE in response to various educational needs. PFHS is a simulation technique that combines PTT with SPs is proposed in the literature as an adjunct to the conventional relatively decontextualised approach to simulation-based clinical skills training, by refocusing clinical skills simulation on the patient and situate SBE in the wider clinical context and thus increase authenticity. The theoretical arguments for the use of PFHS are compelling and there has been increasing adoption of this approach to simulation over the last few years. However as with all new educational techniques, they use should be firmly grounded on both theoretical and empirical evidence. In recognition of this, in the next chapter, I will present a systematic synthesis of the current evidence of the use if PFHS in relation to training and assessment which will provide further background to this thesis and give rise to the key research questions.

CHAPTER 2 - Current Evidence of Patient-Focused Hybrid Simulation in Training and Assessment

2.1 Introduction to Chapter

In the previous chapter, I discussed the key theoretical underpinnings of PFHS, specifically the value of authenticity and professional context, the need to ensure integration of a range of skills for clinical competence, and the need to place the patient at the centre of the clinical simulation and practice. Since its introduction, there has been widespread adoption of PFHS in clinical education, in both the undergraduate and post-graduate sector, across disciplines within medicine and the wider healthcare community, for both teaching and assessment. This widespread adoption perhaps largely due to some of the strong theoretical arguments that support its use. Indeed, there are studies in the current literature which have evaluated the use of PFHS favourably, in terms of acceptability and trainee satisfaction (Girzadas et al., 2009; Higham et al., 2007; Siassakos et al., 2010).

Despite this, the use of PFHS, as with any educational intervention needs to be based not only on theoretical but also empirical evidence. Simulation, although not always, is often more labour intensive and has a higher upfront cost when compared to other methods of teaching and assessment and thus there are demands for evidence to justify its use (Bradley, 2006; Wenk et al., 2009). In addition, any training intervention or new assessment should not only have demonstrable effectiveness, but should not have any untoward side-effects (Bond et al., 2007). Within the wider simulation literature, there have been a number of studies demonstrating a lack of training effect, or even worse unwanted learning effects, such as over confidence or inducing unnecessary stress (Olympio et al., 2003; Quilici et al., 2005). In this chapter, I will present a systematic review of the current evidence of PFHS in the literature.[2]

[2] This systematic review, which contributed to the key research questions, was completed in 2011. Since this systematic review was conducted, 2 additional papers have met criteria.

2.1.1 Assessing quality of research in SBME:

Evaluating the quality of evidence of training interventions required additional considerations to the established levels of evidence-based medicine, which is the standard used to assess the literature for clinical interventions (M. Harden, 1999). A commonly used approach for assessing evidence in medical education literature which is recommended by BEME (Best Evidence in Medical Education) is by using the modified Kirkpatrick hieraracby, a 4-tier system as follows (Hammick et al., 2010):

- Level 1: Reaction – this is a low level evaluation assessing learner's perceived value, acceptability and satisfaction with the training intervention
- Level 2a: Modification of attitudes – This is a measure of changes in attitude, for example are trainees more aware of safety issues after training.
- Level 2b: Modification of Knowledge and Skills – measuring in change in knowledge, i.e. via written tests, or skills usually via simulation.
- Level 3: Behavioural change – evaluation of transfer of learning to the workplace
- Level 4a: Organisation practice – Evaluating effect of training on changes to the organisational delivery of care
- Level 4b: Benefits to patients- Evaluating the effect of training on patient outcome.

Simulation-based assessments should possess a range of qualities to justify its use. In addition to feasibility and acceptability by both assessors and assesses, they should also have the desired psychometric properties, i.e. reliability and validity (Schuwirth and van der Vleuten, 2003). This is particularly important with respect to providing evidence for use in high-stakes examinations.

Reliability is the measure of whether a tool can produce similar results under consistent conditions. For assessment, there are several types as follows (Cohen et al., 2007):

- Inter-rater reliability – this is a measure of the degree to which the results of an assessment conducted by two or more assessors correlated or agree.
- Test-retest reliability – this is a measure of the degree to which test scores are consistent from one administration of the test to the next.

- Internal consistency – this is a measure of the consistency of results of the different items within a test.

Validity is the measure of whether a test measures what it intends to measure. For competence in clinical skills, it is a measure of whether the test measures actual competence. There are several common approaches to measuring validity used in simulation as follows(Downing, 2003):

Face Validity - This is a crude measure usual based on expert opinion of whether a test is appears to measure what it was designed to measure. For simulation-based assessments, this is often defined in the literature as how realistic a simulation is, though it is a common misconception (Carter et al., 2005). Instead it should be related to whether the simulation-based assessment on face value assesses the competency of interest (Crossley et al., 2002) For instance, a banana is arguably less visually realistic than a synthetic skin pad. However if the purpose of the test is to assess basic suturing skills, both are probably equally valid(Kyaw Tun et al., 2011).

Content Validity – This is a measure of the comprehensiveness of an assessment in relation to the desired criterion and usually also based on expert opinion (Cohen et al., 2007, p. 137). For example, a simulation-based assessment intended to assess a doctor's ability to perform a full cardiovascular examination would have limited content validity if it only assessed auscultation skills.

Construct Validity – This measures the extent to which a test conforms to a theoretical construct. One of the most common approaches to determining construct validity in simulation-based assessments is by measuring its ability to discriminate people of different experience levels, on the theoretical assumption that more experienced personnel should perform better (Ansell et al., 2012; Cohen et al., 2007, p. 138).

Concurrent and Predictive Validity – These are types of construct validity, which measure the extent to which a test correlates to another test which had previously been proven to be valid (Cohen et al., 2007, p. 140). Ideally, for assessment of clinical competence, simulation-based assessments should correlate with real workplace performance (Wass et al., 2001). Alternatively concurrent validity can be measured by determining correlation with other previously validated simulation-based assessments. Concurrent validity measures correlation of two tests at the same time, whereas predictive validity measures correlation of two tests administered at different times, e.g. does a simulation-based assessment predict future performance in the workplace.

For simulation-based assessments, there are two aspects of reliability and validity that need to be considered. First the simulation itself needs to be reliable and valid, in that it must reliably and adequately allow assessees' to demonstrate the competencies of interest. Secondly the rating tool that measures assessees' performance in the simulation must also be reliable and valid (Bray et al., 2011).

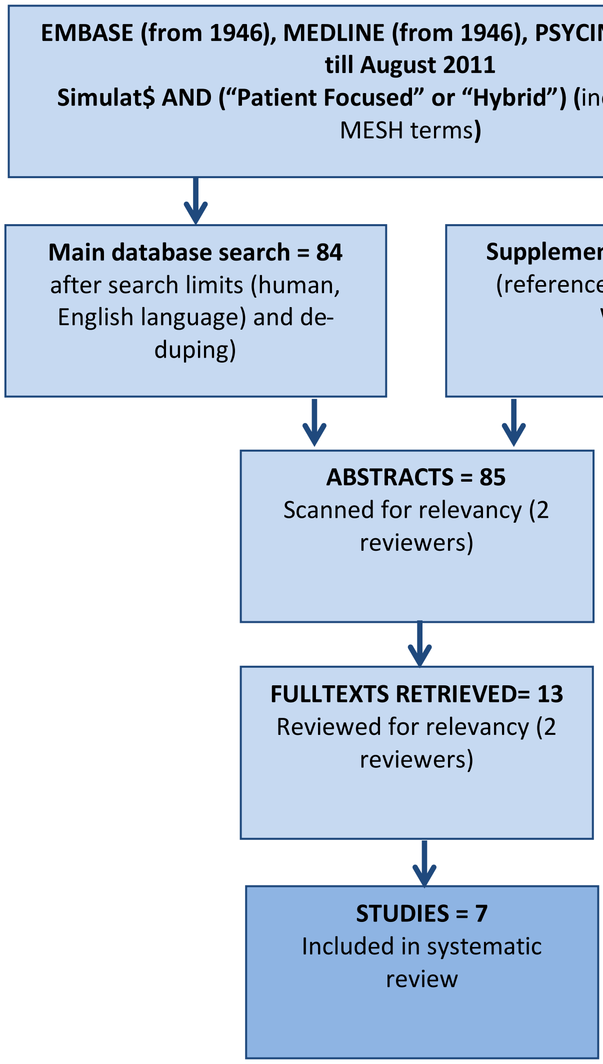
Take for instance, a cardiovascular examination requires a number of competencies including performing physical manoeuvres to identify and illicit clinical signs, as well as communication skills for encourage patient co-operation and to develop trust and rapport. Even if a simulation can allow an assessee to demonstrate all the relevant competencies, if the rating tool only rates the technical or communication aspect, then it may diminish the validity of the assessment. Likewise, if the rating tool used to measure the performance produces inconsistent ratings, then its reliability and thus overall utility is diminish. The aim of this synthesis is to provide an in-depth review and analysis into the current evidence with respect to training effectiveness and psychometric properties of PFHS.

2.2 Methods:

An English language literature search of original articles investigating the use of PFHS for training and assessment of clinical skills, using EMBASE from 1946, MEDLINE from 1980 and PSYCINFO from 1967 till August 2011 was undertaken (Figure 2.1). Due to the varied use of simulation-related terminology in the literature, broad search terms were used. Reference lists of all included full text articles and key review papers were also reviewed for additional papers. Finally, a freetext search using several Internet search engines was performed. Search terms used are presented in Figure 2.1.

Two reviewers (JKT and AG) assessed the retrieved articles independently for relevancy according to pre-determined inclusion and exclusion criteria as below. Disagreements were discussed and reviewed until agreement was reached.

Figure 2.1: Systematic Review Search Strategy



Inclusion Criteria:

- Primary research studies
- Studies evaluating effectiveness of PFHS for training of clinical skills
- Studies objectively investigating the use of PFHS for assessment of clinical skills
- Comparison studies (studies comparing PFHS to other training modalities).

Exclusion Criteria:

- Non full-text and articles.
- Secondary studies, commentaries or opinion papers
- Subjective evaluation of training effectiveness, i.e. Kirkpatrick 1
- Studies that only subjectively evaluate PFHS as an assessment tool, e.g. student and examiner acceptability.
- Poor or unclear study design

Due to the heterogeneous nature of the study methods and outcomes, a meta-analysis was not performed. Instead, a qualitative synthesis of the finding was performed. Articles were reviewed for study aims, design, training and /or assessment methods, competency domain, outcomes and limitations.

2. 3 Results:

A total of 7 studies were identified over the period of 2006 till October 2011. Study participants included medical students, surgical and medical residents, nurses and paramedics. A range of clinical skills was studied including cardiovascular and respiratory examination, basic procedural skills, communication skills and obstetric skills.

2.3.1 Effectiveness of training of clinical skills with PFHS

Three primary studies evaluating the use of PFHS for clinical skills training that met the inclusion criteria were identified. Two studies compared the use of PFHS to other conventional methods of training. A summary of the key findings is presented in Table 2.1. Siassakos et al (2010) reported a randomised controlled trial comparing the use of PFHS (intervention) to small group teaching (control) for obstetric delivery skills. 24 medical students were randomised into either the intervention or control group. All students participated in an initial session combining didactic teaching and simulation session using a part-task trainer in isolation. The intervention group then participated in a PFHS session combining an obstetric delivery part-task trainer and SP with specific attention to communication skills. The control group received further didactic teaching and no additional communication skills training. Following training, students were assessed in a PFHS scenario delivering a baby with shoulder dystocia. Performance was rated by a single blinded expert using the Patient Perception Score (PPS) a rating tool for assessing communication skills that was previously validated in the real clinical setting, but not in the simulated setting. The authors reported that intervention group performed significantly better ($p<0.5$) than the control group following training. However, the authors also acknowledge some confounding factors, including the use of a single rater which may have introduced bias. Another consideration to this study is that although the training was designed for training technical and non-technical skill simultaneously, there was no mention of the technical ability following training.

Simon et al (2012) reported a prospective case-control study comparing PFHS (intervention) to didactic training (control) for heart and lung auscultation skills for paramedics. 22 paramedics participated in the intervention group and 18 in the control group. The intervention group participated in a series of 5 PFHS scenarios incorporating a simulated stethoscope (Ventriloscope®) and SP whilst the control group received audio only auscultation using pre-recorded clinical auscultatory sounds and a didactic lecture. All participants participated in a pre-test, immediate post-test and delayed post-test (4 weeks after training), which consisted of a series of 10 clinical scenarios. Both groups performed significantly better in the immediate post-test in comparison to the pre-test ($p<0.05$). This effect persisted in the delayed post-test within the intervention group, i.e. the delayed post-test performance was significantly higher than the pre-test. However, within the control group, the delayed post-test performance was not significantly different from the pre-test. The authors conclude that the use of this specific PFHS is better than isolated auscultation training using pre-recorded sounds, particularly with respect to longer-term retention of skills. There are, however a number of limitations to the study, the most crucial of which is the lack of validation information of the assessment tool used to measure performance in the pre and post-test.

One study investigated the effect of feedback on effectiveness of PFHS training. Moulton et al (2009) reported a stratified randomised controlled trial comparing PFHS with SP-led feedback to PFHS without feedback for training of clinical/surgical skills. Participants were medical students or junior surgical residents who were first stratified into the respective experience levels and then randomised into the intervention and control groups. All participants completed 4 PFHS scenarios in total, each of which assessed a difference clinical skill as follows: wound closure; urinary catheterisation; application of a cast; and skin lesion removal. After completing 2 scenarios, the intervention group received SP-led feedback specifically with respect to communication issues with the aid of video review of performance, whilst the control group did not. All participants then completed the final 2 PFHS scenarios. Performance was rated by blinded expert raters using previously validated rating tools including a task-specific checklist (TSC) and global rating scale (GRS). The authors reported that in the initial 2 scenarios (pre-training), there was no statistical difference in communication performance between the two groups. Communication skills performance in the final 2 scenarios (post-test) was significantly better in the intervention group in comparison to the control group. No difference was found between the two groups with respect to technical skill performance. A key limitation to this study was the use of different clinical scenarios in the pre and post-test, which may have been a confounding factor.

2.3.2 Current evidence of psychometric properties of PFHS-based assessment

Four studies that investigated the psychometric properties of PFHS were identified which met the inclusion criteria. A summary of the findings is presented in Table 2.2.

Isenberg et al (2011) reported a study evaluating the use of PFHS to assess a range of clinical and basic surgical procedures. 670 medical students over a period of 3 years participated in an OSCE involving a series of 5 PFHS stations as part of their assessment for the surgical student placement. Performance was measured using a range of tools. A task-specific checklist was used to rate technical skills and data collection (e.g. acquisition of patient history). Communication and interpersonal skills were rated using a tool developed in-house and which was the standard assessment tool used in the institute based on Likert scales. Each station was rated by a SP who had been trained to assess students' performance. Cronbach's alpha for the various components of the student's performance was as follows: 0.58 for data gathering, 0.71 for procedural skill, and 0.97 for communication/interpersonal skills. Performance in the OSCEs positively correlated with the students overall scores for their clinical placement, which broadly measures their clinical ability in the workplace.

However the authors do not specify what the clinical placement scores actually measure. The authors conclude that the OSCEs demonstrated construct validity.

LeBlanc et al (2009) reported a quasi-experimental study, evaluating the use of HS for assessment of clinical and basic surgical skills competence. Sixteen medical students and 16 junior surgical residents were invited to complete a 4 station OSCE involving a series of PFHS scenarios. Performance was rated by a single expert assessor per station using a previously validated TSC and GRS for technical skills and a communication skills score devised and previously utilised by the institute. Intraclass reliability was satisfactory ranging from 0.61 to 0.75. Residents performed significantly better than medical students in terms of technical skills as measured by both the TSC and GRS. Residents performed better than medical students in 1 of the 5 items in the communications skills score (coherence in communication), but no difference was found in the other items. The authors suggest that that this may be due to current lack of communication skills training amongst residents.

Verma et al (2011) reported a study investigating the use of PFHS combining a simulated stethoscope (Ventriloscope®) and SP for respiratory examination as part of a 12 station OSCE. 285 medical students participated in the study. Performance was measured using an OSCE checklist typically used at the institute, rated by one examiner. Scores in the PFHS station correlated significantly with scores from other stations.

Ponton et al (2011) evaluated the psychometric properties of a 7 station OSCE of which two stations consisted of PFHS scenarios. 14 surgical residents participated in the study. Performance was rated with a TSC and GRS by one examiner. Intraclass reliability was highly variable when performance was scored with the TSC alone (a = 0.19 to 0.86), though this improved when TSC scores and GRS were combined (a = 0.87 to 0.93). Overall performance in the 7 stations of more experienced trainees was significantly higher than in less experienced trainees, though there was no sub analysis of the specific PFHS stations to provide validity data.

Table 2.1: Summary of Studies Investigating Training Effectiveness of PFHS

Author	Study aims	Study Design /KP level	Participants	Training intervention	Outcome measures	Outcomes	Study limitations
Siassakos et al (2010)	Compare PFHS training to small group tutorial for obstetric	RCT KP 2b	Medical students (N=24) randomised to Intervention Group (n= 11) Control Group (n=9) due to dropout	Intervention group: 1 session didactic training and PTT training. 1 session PFHS training (SP with obstetric simulator) with communication skills Control group: 2 session didactic training and PTT training. No communication skills training.	Post-training simulation-based assessment (simulated delivery with shoulder dystocia). Performance rated with PPS Score (measuring communication skills) by one expert rater	Intervention group significantly better than control group ($p<0.05$)	Single rater. Lack of demographics of participants.
Simon et al (2011)	Compare PFHS to sound library for training heart and lung auscultation skills	Prospective case-control study. KP 2b	Paramedics (N=40): Randomised to Intervention (n=22), Control (n=18)	Intervention group: Series of 5 PFHS scenarios (Ventriloscope (R) and SP) Control group: Audio only auscultation training using pre-recorded clinical auscultatory sounds and a didactic lecture	Immediate and delayed (4 weeks) Post-training simulation-based assessment using non-validated rating tool by one assessor.	Performance of both groups significantly improved in the immediate post-test. Performance significantly improved in delayed post-test in comparison to pre-test in intervention group. No significant difference between delayed post-test and pre-test performance for control group.	Possible selection bias. Non-validated rating tool. Single assessor
Moulton et al (2010)	Compare PFHS training with SP-led feedback to without for clinical skills	Stratified RCT KP 2b	Medical students (n=16) and Junior surgical residents (n-16). Stratified to training level and randomised into Intervention (n-16), Control (n=16)	Intervention: 2 PFHS followed by SP-led feedback on communication, then 2 PFHS Control: 2 PFHS followed by no feedback, then 2 PFHS	Post-training simulation-based assessment (2 different clinical scenarios) Rating using task-specific checklist and Global rating scale, by two expert raters	Intervention group significantly better than control ($p<0.05$). No difference between groups in technical performance ($p>0.05$).	Different scenarios assessing different competencies pre and post training.

Table 2.1: Summary of Studies Investigating PFHS for Assessment

Author	Study Aims	Study Design	Participants	Assessment format	Performance measures	Outcomes	Study Limitations
Isenberg et al (2011)	To evaluate PFHS-based assessments of clinical skills in surgical clerkships	Prospective cohort study	670 medical students in series over 3 years	5 PFHS stations as part of OSCE	OSCE checklist rated by single trained SP	Internal consistency range from 0.58 – 0.97. PFHS scores correlates with surgical clerkship rand other specialty clerkship ratings, i.e. construct validity demonstrated	Clerkship assessment not defined
Verma et al (2011)	To evaluate PFHS based assessment of respiratory examination skills.	Prospective cohort study	385 medical students	12 station OSCE, of which one station is a PFHS of interest (respiratory examination).	Institute's OSCE rating tool. One examiner per station	Strong correlation between PFHS performance with performance in other conventional OSCE stations (Spearman's correlation coefficient – 0.97 $p<0.001$)	No direct correlation between PFHS and patient
LeBlanc et al (2009)	Evaluate use of PFHS for basic surgical/ clinical skills assessment	Quasi-experimental	Medical students (n=16) Junior surgical residents (n=16)	OSCE with 4 PFHS stations.	Performance rated with previously validated tools: GRS, TSC for technical skills; Communications skills rating tool. 1 examiner per station	IRR acceptable for : GRS (ICC = 0.76, a= 0.75); TSC (ICC = 0.73, a = 0.65). Communication skills (ICC = 0.61, a = 0.81) Residents had significantly higher scores than students for GRS and TSC scores students ($p<0.05$) and one communication score item ($p<0.05$). No difference in the remaining items No correlation between communication and technical skills	
Ponton-Carss (2011)	Evaluate psychometric properties of an OSCE	Prospective cohort study	14 surgical residents	7 station OSCE (2 stations of which are PFHS)	Rated using TSC and GRS by one rater	IRR of the TSC and GRS scores PFHS stations were acceptable for technical performance (0.19 to 0.86)	No sub analyses of PFHS data

2.4 Discussion:

The aim of this synthesis was to evaluate the current evidence of PFHS in the clinical skills education relation to objective training effectiveness and assessment properties. The current evidence base on the use of PFHS is limited, though this is unsurprising given the relative novelty of this technique, which has only been popularised and adopted into mainstream medical education in the last decade. There is however a range of clinical skills that that have been simulated using this method in the literature which reflects the potential of this technique.

From the perspective of clinical skills training, there is at current limited evidence demonstrating training effectiveness. The available evidence shows that there is in general a positive learning effect with the use of PFHS for training for a range of clinical skills. Specifically, training effectiveness of PFHS has been demonstrated in both the domains of technical (procedural) and communication skills. In addition to demonstrating effectiveness of PFHS, one study demonstrated the increased effectiveness of PFHS with feedback. Whilst the available evidence is promising, it should be interpreted with care. First, these studies demonstrate improvement of skills in the simulation setting, i.e. at Kirkpatrick level 2b. This is problematic as it is difficult to determine whether the learning effect is due to improved “simulation competence”, i.e. due to familiarisation of the simulation, or whether actual clinical competence has improved (Bradley and Postlethwaite, 2003a). Studies have indeed demonstrated significant learning effects even after between two simulations performed in series (Gaca et al., 2007).

The lack of evidence demonstrating transfer of training from PFHS to the real workplace is unsurprising given the ethical considerations of conducting such research, which is a recognised issue when researching simulation training in general (Mcgaghie et al., 2010). There have been a number of papers calling for more studies measuring transfer of training to the clinical workplace with randomised controlled trials comparing the effect of simulation to no simulation training on performance in the real clinical setting. However, this can be difficult to justify particularly with respect to invasive procedures where patients may come to harm. Consider for instance the ethical dilemmas that may be encountered in conducting a hypothetical study comparing a group of novice foundation year doctors (interns) with no prior experience in central venous catheter insertion are randomised to receive simulation or no-simulation training for inserting central venous catheters and then subsequently assessed for competence on a real patient, without supervision (to control for confounding factors). Given the limited but available evidence on simulation-based training, such a study would be difficult to justify on ethical grounds.

From the perspective of assessment, current studies demonstrated that reliability is variable. The reasons for this are not certain, but likely to be multifactorial. One study demonstrated a variation in inter-rater reliability depending on the type of assessment tool applied (Ponton-Carss et al., 2011). Another reason may be that PFHS measures integrated performance, which may have greater scope for a varied, but yet still appropriate approaches to the same clinical problem, which may result in more variation between different assessor's interpretations of assessee's performance and diminish reliability.

With respect to assessment validity, there is some limited evidence beyond face validity. One study demonstrated a correlation between PFHS assessment scores and clinical placement scores, and another study showed a correlation between PFHS assessment scores and scores in other conventional OSCE stations, thereby demonstrating a degree of construct validity (Verma et al., 2011). There is also some evidence of discriminant validity, though this is only demonstrated predominantly in the domain of technical procedural skills, with minimal evidence with respect to non-technical skills. This is lack of evidence is problematic. A key purpose of using PFHS is to allow assessment of non-technical ability such as communication skills. The reasons for this lack of evidence are uncertain. In one of the studies reviewed, the authors suggested a possible reason that PFHS assessment could not differentiate their relative experts (surgical residents) and novices (medical students) and in terms of communication skills and hence did not demonstrate construct validity, was a lack of communication skills training during residency (LeBlanc et al., 2009). Another reason for this finding may be due to the challenging nature of measuring competence in communication skills which is ill defined (Marsden, 2014). Specifically, the conventional approach of assessment of technical competence, may not apply to the measurement of non-technical skills, which is considered to be relatively difficult to measure (Lane and Rollnick, 2007).

There are a number of limitations to this review. First, although broad search criteria to ensure all relevant articles according to the inclusion criteria set were included, grey literature was not reviewed which may provide further evidence of the use of PFHS. In addition, from my experience as a teacher and assessor, there are examples of good practice that are unreported that may provide insight into this field. Finally, whilst the focus of this review is of PFHS, there may be evidence from the literature of other simulation techniques that may be transferable to this context. In particular there is limited by existing evidence from the wider simulation literature demonstrating transfer if training to real clinical practice (Lynagh et al., 2007).

With these limitations in mind, it can be seen from this literature review that despite the strong theoretical reasons for use of PFHS and its documented use for almost a decade, there remain several areas that require further research. In addition to a current limited evidence base, there is no high level evaluation of training effectiveness (Kirpatrick 3 and 4). Specifically there requires further research comparing training and assessment with PFHS to with real patients.

Another area that requires further research in PFHS is in its value from the perspective of simulating professional context. In addition, although a key argument for using PFHS is to introduce context to simulate complexity and risk, there is yet limited research in this area. A previous study by Higham et al (2007) investigating the use of PFHS to simulate gynaecological examination scenarios of different levels of patient complexity demonstrated a subjective perceived increase in challenge in the more complex scenarios. However, there is no study objectively assessing the effect of simulating clinical complexity with PFHS on performance.

Finally, within the current literature of PFHS and wider literature on SBE in general, there is very little empirical data that informs us of the process of teaching and learning in simulation. Indeed it is well-recognised that much of the research in the general educational literature focuses on end points in terms of learning outcomes, but there is relatively little that is known about the finer details of happens in a classroom. Yet the classroom is an understudied, complex environment which may provide a multitude of factors that can influence learning outcomes and has therefore been termed as the "black box" (Black and Wiliam, 1998) . Equally in the studies evaluating PFHS reviewed in this synthesis and within the wider literature that was scoped for this synthesis, there is little information about how learning took place and what the role of the tutors were. In addition, within the current SBE literature, conclusions often posit the research outcomes to be directly attributable to a simulation activity itself. However in reality, simulation is used in conjunction with other educational techniques such as lectures, small group tutorials and conducted under the guidance of a tutor, who themselves may have variable attributes and qualities.

In view of these findings, the key areas and questions that I aim to investigate in the empirical component of this thesis are as follow:

1) The role of context in simulation-based assessments
A) Can clinical challenge in simulation-based assessment be simulated through modifying clinical context with PFHS
B) How do clinicians perceive and compare simulation-based assessment with PFHS to PTT used in isolation.

2) Comparing the Use of PFHS to Real Patients for training and assessment.
A) How do students who train with PFHS compare to those who train with real patients
B) How does assessment of clinical skills with PFHS compare to assessment with real patients
C) What are learners' experience of training with PFHS and how does this compare to when performing on real patients.

2.5 Summary

In this chapter, I presented a synthesis of the current evidence regarding the use of PFHS in clinical skills education with respect to training effectiveness and assessment with a focus on psychometric properties. The current evidence base is limited, but promising. With respect to training effectiveness, positive learning outcomes have been demonstrated, though post-training learning effect has been demonstrated using simulations rather than performance in the real clinical workplace. For assessment, there is some evidence of construct validity, whereby clinical competence measured with PFHS correlates to other accepted assessment measures, but no current research comparing assessment with PFHS to with real patients. In light of these issues, I proposed a series of key research questions that I aim to inquire in this thesis. In the next chapter, I will present some methodology considerations. Specifically, I will discuss the underlying philosophical and epistemological considerations in researching SBE and in relation to the key questions proposed.

CHAPTER 3 Methodological Approach and Philosophical Considerations

3.1 Introduction to Chapter

In this section, I will discuss the rationale for my methodological approach to the empirical component of my inquiry into PFHS. I will draw on epistemological and ontological considerations, on the basis that these in turn give rise to methodological assumptions (Kneebone, 2002; Krauss, 2005). I present a critical overview of the qualitative and quantitative paradigms and present the benefits and limitations of each with respect to research in simulation-based medical education due to its complex nature. Finally I present my rationale for using a mixed-method approach to my inquiry.

3.2 Qualitative and Quantitative paradigms

In this section, I will present an overview of the qualitative and quantitative research paradigms as a basis for subsequent discussion into their suitability for simulation based research. There are a number of other major research paradigms used in education research such as the critical theory approach, though they are not aligned with the aims of this thesis and therefore will not be discussed (Chapter 1, Cohen et al., 2007).

3.2.1 The Quantitative Paradigm

The quantitative research approach is based on the positivistic paradigm. Positivism is derived from the ontological position of realism - that truth and reality are singular (Feilzer, 2010). In this respect, reality is absolute and exists independently from human perception (Sale et al., 2002). In simple terms an object, for example an apple is always an apple regardless of the observer. From this perspective, research, inquiry and the search for the truth requires objectivity, where the investigator aims to study a given phenomena without influencing or being influenced by it (Schrag, 1992). This positivistic philosophy is the basis of much of the research in the natural and biomedical sciences (Tavakol, 2009). Methods and techniques employed for this approach to inquiry include control, randomisation, blinding, and highly structured protocols, to minimise biases of observation and other confounding factors. The purpose of this methodological approach is to find generalisable truths and "laws" about a phenomena (Tavakol, 2009).

Quantitative research is aims to assess the relationship between cause and effect through restricting confounding variables. Quantitative research also typically requires large samples to achieve sufficient statistical power such that research findings can be generalisable. The quantifiable data generated therefore is often perceived to have higher credibility particularly with policy makers (Johnson and Onwuegbuzie, 2004).

Quantitative research however, due to its nature, often does not take into account local contextual factors. As an underlying key principle of quantitative research is to remove confounding variables, context in which phenomena are often seen as "white noise". Yet, the context in which an effect occurs is arguably context dependant, and as researchers may miss out on important phenomena. From a practical prospect, knowledge generated may also be too abstract for direct application to specific situations, context and individuals. A classic example of where the lack of consideration of context yielded questionable findings and conclusions is in the initial studies measuring the intelligent quotient (IQ) of persons of different race and ethnicity (Marks, 2010). In these studies, IQ was considered to be an absolute measure of intelligence, though it is now widely acknowledged that it needs to be interpreted in like of a host of factors including social background, country of origin and so forth.

The lack of consideration of context can therefore lead to a commonly encountered situation of a theory-practice or policy-practice gap, whereby the generalizable theories derived from quantitative research do not translate into messy world of real practice (Cohen et al., 2007, p. 19) . A final criticism that has been suggested is that as quantitative research aims to test hypotheses, it is at risk of confirmation bias (Johnson and Onwuegbuzie, 2004). By this I mean research becomes a means to prove or disprove what one already suspects as opposed to finding out the "truth". This can manifest itself in a number of ways, first in terms of conducting the research itself in terms of study protocol, but also in terms of interpretation of results. As stated earlier in Chapter 2 in the systematic review, whilst a number of the studies demonstrated mixed outcomes, some with "negative" outcomes, overall, the studies were concluded "positively" in favour of the use of PFHS.

3.2.2 The Qualitative Paradigm

The qualitative approach to inquiry is derived from the interpretivistic and constructivist paradigms (Schwandt, 1994). From an ontological perspective, this paradigm is based on the assumption that there are multiple realities and truths based on an individual's construction of reality (Morgan and Smircich, 1980). As reality is socially constructed, it is not fixed or finite. Reality therefore cannot exist without participants and researchers and can only exist through ones' perception. Methodological approaches employed within this paradigm aim to elicit truths about a phenomena through the lens of an observer and or participant. As such, techniques such as interviews and focus groups, which seek to obtain participants' perspective, and ethnography where phenomena are understood through the observers' perspective, are utilised. At the same time, study samples do not necessarily represent large populations. Instead, small focused sampling is used to provide more depth and richness to inquiry.

There are a number of key strengths to qualitative research. First, it allows researchers draw on their own personal experiences of phenomena as well as those of participants. As such, phenomena can be interpreted through an individuals construct and understanding of the world. Data is also collected in a naturalistic setting without the strict methods of reducing variables as in quantitative research (Black, 1994). Therefore it allows inquiry of phenomenon in its usual context. This is particularly useful for describing and understanding complex phenomena. Finally qualitative research is usually dynamic, therefore allowing researchers to conduct more in-depth research through modifying their angle of inquiry as they generate new findings.

There are however also a number of limitations to qualitative research. Due to the lack of control for variables and small sample size, it is difficult to make quantitative predictions and generalisations. Data collected is more easily influenced by researchers personal biases and idiosyncrasies. As such it is often considered not to be 'hard' evidence by policy and decision makers ((Johnson and Onwuegbuzie, 2004). Both qualitative and quantitative paradigms have benefits and limitations in educational research. In the next section, I will discuss some considerations in the research of simulation-based medical education and benefits and limitations of each approach.

3.3 Establishing a Research Approach to Simulation-Based Medical Education

The choice of which research paradigm to use must first and foremost be appropriate for the area of inquiry to yield valid inferences. However, research study design needs to also take into account other factors such as resources available. In addition, research should not be considered as a means to an end. It needs to take into account the purpose of research and who the stakeholders are, i.e. research sponsors, policy makers and peers in the research field (Cohen et al., 2007, p. 80).

There are a number of issues that need be taken into account when researching simulation-based education. Simulation is a complex educational intervention with a multitude of potential confounding variables (McGaghie et al., 2011). As discussed earlier in Chapter 1, simulation encompasses a wide range of activities and can vary in terms of comprehensiveness (whole-task/part-task) and fidelity. In addition, the effectiveness of a simulation activity for training and assessment depends not only the simulation itself, but other factors such as how it fits into the wider training curriculum, other co-existing educational activities such as workplace learning, and crucially the individual students and clinicians as participants themselves. It is widely acknowledged that there are many different types of learners, some favouring one approach over another (Pashler et al., 2008).

Simulation is an activity that requires participants to be involved from within, and therefore a participant's personal construct of their experience in simulation is likely to influence its effectiveness. For example, if we want to assess a participant's real life performance and ability through simulation, there may be a need to ensure that the participant 'buys into' the simulation and suspends disbelief so that they can perform as they do in real life (Seropian 2003). However this is likely to vary from one participant to another. From this perspective, a qualitative approach can offer some benefits. By seeking to understand participants' perspective of simulation activities, underlying mechanisms for success and failure may be elicited.

However there are also some key limitations to qualitative research in this context. First, in this thesis, I am specifically concerned with the suitability of PFHS simulations for training and assessment. With respect to assessment of competence if simulation is to be used for high-stakes examinations, quantifiable evidence with statistical significance to support its use is essential (Scalese et al., 2008). From the perspective of training, any new intervention should be robustly evaluated for effectiveness, or worse, negative learning effects (Bond et al., 2007). In order to make decisions on whether to use simulation, quantifiable evidence should be therefore available. For policy makers, this is particularly important in medical education where the predominant paradigm is positivistic. Qualitative research may not provide the type of evidence that is seen as credible by stakeholders and as such is limited. There some authors who have highlighted the need for more 'hard' evidence in the form of randomised controlled trials (Weller, 2004).

Based on these issues raised, I argue that, a purely quantitative or qualitative approach in isolation has limited scope for investigating simulation-based medical education. Instead, I propose the use of a mixed-methods approach.

3.4 The Mixed-Methods Approach

Mixed-methods research can be defined as 'the class of research which mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study' (Johnson and Onwuegbuzie, 2004). The mixed-methods approach utilises elements of both qualitative and quantitative paradigms in a complimentary manner to offer additional insight into an area of inquiry. It is not a simple haphazard marriage of the approaches and all elements of the research need to be considered holistically.

Although it is a relatively new paradigm, mixed methods has in fact been practiced for some time but perhaps had not been labelled as such (Johnson et al., 2007). The relative novelty of this paradigm and rapid adoption has resulted in a wide range of approaches to applying this paradigm in research (Schifferdecker and Reed, 2009). There are however some key common philosophical assumptions that underpin the use of mixed-methods. First, it is based on the premise that all methods (including those belonging to other research paradigms such as critical theory) can be classified into qualitative and quantitative research. Secondly, it posits that both paradigms can and should co-exist in a single area of inquiry. Thirdly, the mixed-method approach is grounded in the philosophical stance of pragmatism, which recognises the absolute positivisitic perspective of the world yet acknowledging the importance of context and situation. Johnson & Onwuegbuzie (2004) summarised the main purposes for using the mixed-methods approach in research as follows:

- *Triangulation* - in order to seek convergence, and verification of findings.
- *Complementarily* - through the use of different approaches, greater clarification of findings can be achieved.
- *Initiation* - paradoxes and contradictions can be found through a more holistic approach to research. Modification and new hypothesis generation through the process may be tested.
- *Development* - through use of the findings form one method to help inform another method.
- *Expansion* through increasing breath and depth of inquiry by using different approaches to inquiry.

Mixed-methods research can be conducted in a number of ways. Qualitative and quantitative research can be conducted concurrently (in parallel) to produce research that has breath and depth as and for cross-validation between findings from the two arms of research (Leech and Onwuegbuzie, 2009). It can also be conducted sequentially to allow data from one part of a study to guide research in another. For example, qualitative research such as interviews can be used to follow-up some quantitative research in order to better understand the “numbers”.

There are a number of criticisms against the use of mixed-methods. Many purists may argue that qualitative and quantitative paradigms are incompatible and cannot be mixed (Howe, 1988). However, some authors including myself would argue that even the purist quantitative research has an element of subjectivity in terms of how data is ultimately interpreted and how inferences are drawn, according to the researchers belief and understanding of the world (Johnson and Onwuegbuzie, 2004; Sale et al., 2002). Likewise, objectivity is often introduced into qualitative research whereby data is coded, quantified and sometimes subject to statistical testing. Therefore in some respect mixed-methods often takes place whereby quantitative and qualitative approaches are used in combination during inquiry even though it is not explicitly labelled as such. Even for clinicians, the process of formulating different diagnoses of patients, which in itself can be considered a form of inquiry usually requires a mixed methods by combining more subjective elements such as history taking with more objective elements such as diagnostic tests (Schifferdecker and Reed, 2009).

Another criticism with mix-methods is the large variation of interpretations as discussed above resulting in inconsistencies in methodological approach. There are no strict formulas for conducting mixed-methods research, though this in itself should not be considered to be a criticism. The open nature of mixed-methods also allows researchers to gain. A further issue with mixed-methods is the practical limitations in terms of cost and resources. Due to the use of two or more research approaches, it is particularly labour intensive. Whilst conducting this type of research it is important that such constraints do not dilute the quality of the research (Johnson and Onwuegbuzie, 2004).

I will now revisit the key areas of empirical research of my thesis, and discuss how the mixed-methods approach will be applied to my research in relation to the key research questions proposed in Section 2.4, which I have presented again below.

1) The role of context in simulation-based assessments

A) Can clinical challenge in simulation-based assessment be simulated through modifying clinical context with PFHS

B) How do clinicians perceive and compare simulation-based assessment with PFHS to PTT used in isolation.

2) Comparing the Use of PFHS to Real Patients for training and assessment.

A) How do students who train with PFHS compare to those who train with real patients

B) How does assessment of clinical skills with PFHS compare to assessment with real patients

C) What are learners experience and process of training with PFHS and how does this compare to when performing on real patients

For key question 1A, I am interested in seeing if clinical challenge can be simulated in procedural skills assessment with PFHS. Although there is some previous evidence of subjective increase in challenge, this needs to be evaluated objectively. Specifically, I am interested in the effect of the simulations' design in terms of level of challenge on participants' performance. Therefore a predominately quantitative approach is used for this component of the study.

For key question 1B, I am primarily interested in the participants' perceptions towards assessment of competence with PTT and PFHS. To explore this, I have chosen to use a quantitative approach (questionnaires with Likert-type questions) to quantify and compare aspects of participants' perceptions towards to the two simulation approaches. In addition, I use qualitative approaches synchronously to explore participant's perceptions in more detail with the aim of triangulation with the semi-quantitative component.

In order answer key research questions 2A and 2B, I aim to compare the training and assessment of clinical skills with PFHS to that with real patients. For this purpose, I will need to employ a predominantly quantitative approach to measure effect of training on learners' knowledge and performance and correlation between performance in assessments with PFHS and real patients objectively.

Finally, for question 2C, I am interested in the learners' experience of training with PFHS and the underlying processes of learning. For this I employed a qualitative approach to gain more in-depth insight into learners' experience primarily through focus group and observations.

3.5 Summary

In summary, in this chapter I presented ontological and epistemological underpinnings of qualitative and quantitative research approaches. I argue that in view of my key research questions and given the complexity of simulation based medical education, that neither qualitative nor quantitative approach to inquiry is sufficient. I therefore proposed a mixed-methods approach and present my broad research strategy to the empirical components of this thesis. More specific details of the research methodology and methods for each study will be presented in the respective chapters. In the following chapters, I will present the research into these areas in detail.

Chapter 4 - Simulating Clinical Challenge with PFHS

4.1 Introduction to Chapter

In this chapter, I will present the first part of a larger overall empirical study of my thesis, which aims to investigate key research area 1, i.e. the role of context in simulation-based assessment of clinical skills through the use of PFHS. The overall study consists of two parts conducted concurrently with the same participants to investigate key research questions 1A) Can clinical challenge in simulation-based assessments be simulated through modifying clinical context with PFHS which is presented in this chapter; and 1B) How do clinicians perceive and compare the use of PTT and PFHS for assessment of clinical competence. The latter will be presented in Chapter 5. An overall map of the study is presented in Figure 4.1. The specifics of the methods and findings will be presented in each chapter respectively.

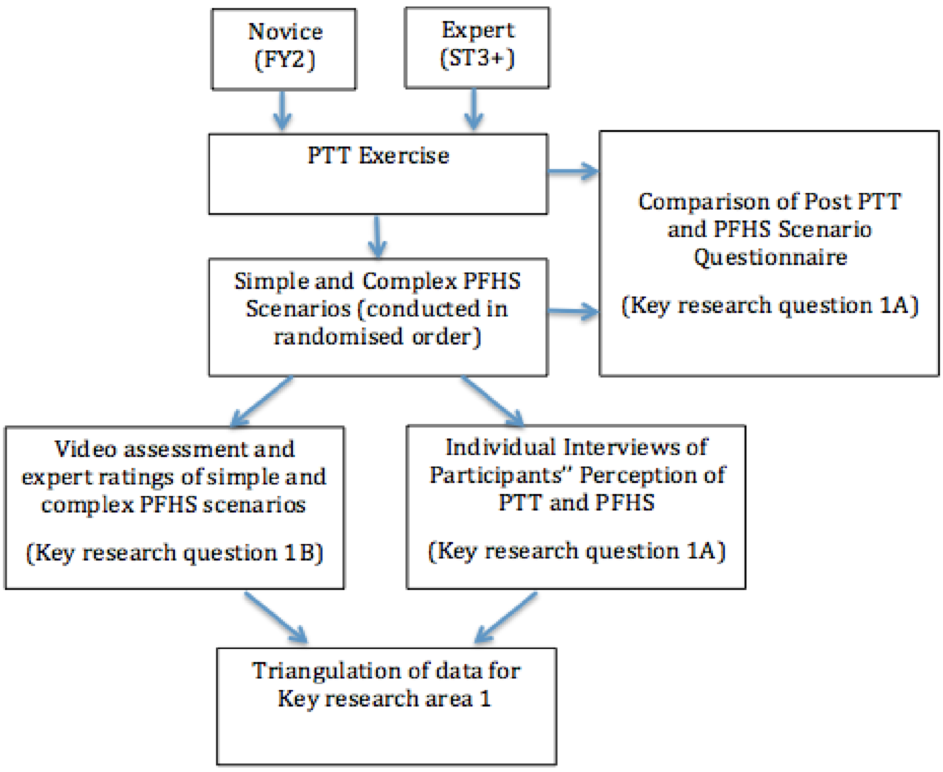


Figure 4.1 Map of Study in Relation to Key Questions 1A and 1B

In this component of the study, I will explore the use of PFHS to simulate varying levels of clinical complexity. As discussed in Chapter 1, clinical competence of any given clinical or procedural skill requires the seamless integrated performance of a range of technical and non-technical skills. For example, when examining a woman's breast, a clinician needs to be proficient from a psychomotor perspective whilst exhibiting the professional behaviour and sensitivity towards the patient; when inserting a central line into an unstable patient, the clinician needs to also have the necessary situational awareness to be aware of any signs of physiological deterioration; when communicating to colleagues from other clinical specialties such as when referring a patient, in addition to good communication, one must have the medical knowledge and expertise with relevant to the clinical case.

In addition to integrated performance, competence requires being able to perform clinical tasks and procedures over a range of clinical complexity (Kneebone et al., 2007). As discussed earlier in Chapter 2, any given clinical task may vary in complexity due to a number or reasons. For procedural skills, the psychomotor challenge can vary between different procedures - inserting a central venous catheter is arguably more complex than inserting a peripheral cannula. For any given procedure, the context may also provide challenge (Kyaw Tun and Kneebone, 2011). From the perspective of psychomotor challenge, cannulating a health patient with "good" veins is easier than performing the procedure on one with "poor" veins such as in an intravenous drug user. The non-technical components may also vary in level of complexity. Using venous cannulation again as an example, cannulating a patient who is cooperative is less challenging than on one who is not.

In this respect, the conventional decontextualised approach to simulation focusing on isolated technical skills is limited in representing the complexities of real clinical practice. As discussed, it has been argued that PFHS may offer an approach to tailoring clinical complexity by means of changing clinical context. However there is no empirical evidence that demonstrates objectively that PFHS can simulate clinical challenge. In addition, although previous studies have explored PFHS scenarios to introduce complexity to clinical skills simulation, there is little in the literature that describes the actual systematic design of the simulation scenarios to the desired level of complexity.

In view of this, the aims of this study are to 1) systematically design of PFHS scenarios to desired level of complexity, and 2) investigate whether clinical challenge can be simulated by objectively measuring the effect of PFHS scenarios designed to simulate different levels of clinical complexity on clinicians' performance.

In order to investigate these, I have chosen to base the simulation scenarios on the management of traumatic wound lacerations within an Emergency Department. There are several reasons why I chose to study the simulation of this particular clinical encounter. First, there is a recognised clinical and educational need to ensure competence in this domain. Management of traumatic skin laceration is a common clinical scenario encountered by emergency medicine doctors, surgeons and emergency medicine practitioners (EMPs) in the Emergency Department. As such, these professional are expected to conduct this competently as documented within their training curriculums such as those written by the relevant training bodies such as a the College of Emergency Medicine and Royal College of Surgeons, UK. Secondly, there were practical reasons for choosing this particular clinical scenario. Previously our research group had created a more basic PFHS simulation for management of traumatic wound lacerations which had been shown to be feasible (providing a basis for this study (Kneebone et al., 2006).

4.2 Methods and Materials

4.2.1 Study Design and Setting

This is a quasi-experimental study to explore novice and expert clinicians' performance on simulations of the same procedural skill across different levels of clinical challenge. This study was conducted at a clinical skills laboratory of a large teaching hospital in London, United Kingdom. Ethics approval was obtained through Brent Ethics Committee (05/Q4408/70).

4.2.2 Selection of Participants:

Participants were doctors across a range of clinical experience from foundation year trainees to specialist trainees in Emergency Medicine. Participants were invited via email and telephone communication from across one UK postgraduate deanery for medical training. 20 participants were recruited in total using quota sampling. All participants were provided with written informed consent, and participated on a voluntary basis. A copy of the consent form is available in Appendix 1. Clinicians were excluded if they had no previous experience of simulation for training and assessment, or no prior demonstration of suturing ability in either simulation or in clinical practice. Participants were stratified into relative novice and expert groups for comparison of validation and performance ratings.

For the purposes of the study differentiation between experts and novices was based on number of procedures performed as opposed to year of training alone. Experts were defined as having worked in EM department and performed over 30 wound closures for traumatic skin lacerations in the Emergency Department.

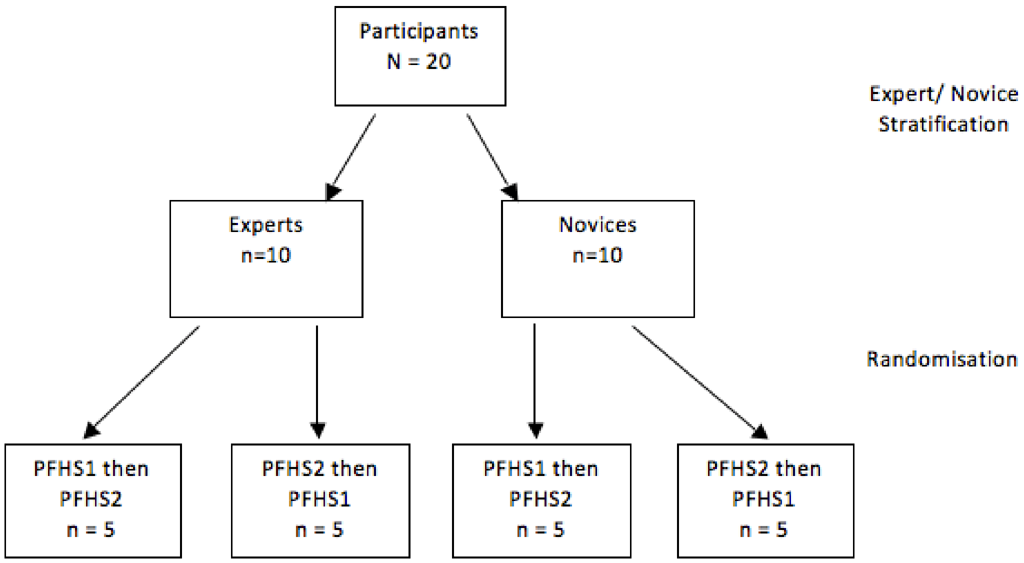
4.2.3 Data Collection

Participants had no prior knowledge of the content of the simulation scenarios before participating in the study. All participants were required to complete a baseline PTT-based scenario based on the management of a traumatic wound laceration to familiarise themselves with 1) style of the briefing instructions for the simulations and 2) the clinical equipment that was provided for the simulations given that there may be slight variation in the equipment used in their normal workplace.[3] Following the PTT-based scenario, all participants were required to complete two PFHS scenarios, one which was designed to be more clinically complex and the other simple as described below. In order to minimise order and learning effects, a counterbalanced design randomised participants to perform either the simple scenario (PFHS1) or complex scenario (PFHS2) first (Figure 4.2). In addition, all scenarios were performed in one session on the same day to minimise external influences on learning and performance in between scenarios, e.g. experience gained in the workplace. All participants were given a standardised briefing for each scenario (Appendix 3).

---

[3] Participation of the PTT-based scenario was also required for the purposes of the second component of this study presented in Chapter 5

Figure 4.2: Stratification of study participants



4.2.4 Simulation scenario design

4.2.4.1 Part-Task Trainer Scenario

This was based on a suturing scenario typically utilised OSCEs. Participants were instructed to clean, anaesthetise and close the simulated PTT wound placed on a bench top. A range of clinical equipment for the exercise was laid out on the table for participants. The scenario was conducted within a clinical skills laboratory environment.

4.2.4.2 Hybrid Simulation Scenarios

The two HS scenarios of different levels of challenge were designed to reflect the real-world practice of managing a patient with a minor traumatic skin laceration that required suturing in an Accident and Emergency (A&E) department. The simple and complex scenarios (Box 1) were constructed to reflect what a Foundation Year 2 doctor (intern) and EM Specialty Trainee Year 3 (3rd year resident) would be capable of performing. The scenarios were developed based on clinical experience of typical cases encountered in an ED by using a process of cognitive task analysis with clinical experts to represent authentic clinical practice.

Most accounts of clinical procedures in textbooks are presented as an idealised set of steps, which does not necessarily reflect the realities of a clinical encounter. Moreover, much of what is performed by experts in clinical practice is automated and tacit, often lying beyond conscious awareness. Cognitive task analysis identifies and documents cognitive processes and mental demands required for performing a task, using a rigorous approach developed from systems engineering. By capturing the finer details of a clinician's approach to a clinical encounter (in this case, the management of an emergency medicine patient with a skin laceration), simulations can be designed that reflect the complexities of actual practice and avoid presenting an oversimplified picture. For performing cognitive task analysis, I have used the approach described by Grunwald et al (2004).

This was an iterative process facilitated by myself, drawing on the experience of S.N. (a senior post-fellowship emergency medicine registrar at time of study) and R.B. (a senior emergency medicine consultant). Scenarios were based on a core framework, which were further developed through serial mental recall and "think aloud" of clinical encounters by S.N. and R.B. This provided information on the skills required, decision-making processes, equipment used, and acceptable variations in approaches to the clinical encounter. Consensus on the content of the simulation scenarios was reached by repeated discussion.

BOX 1: Simple and complex hybrid scenarios

Simple Scenario (PFHS1)

30-year old female pleasant and compliant city worker previously fit and well, sustained laceration to her anterior mid right thigh following a fall from her bicycle at low velocity. No other injuries, or past medical history. Stable on admission.

Complex Scenario (PFHS2)

50 year old intoxicated and disruptive city banker, with known hypertension and penicillin allergy but otherwise fit and well, sustained a laceration to the medial aspect to his mid right upper arm due to a fall over a metal railing. No other injuries or significant history. Stable on admission.

4.2.4.3 Development of Simulated Patient with Simulated Skin Laceration

Wound prosthetic development

Early patient-focused hybrid simulations utilised existing PTT that were designed to be used in isolation on a bench top. In order to achieve realism, these required clever “camouflage” with the use of drapes in order to make the PTT appear to be part of the SP. However, as wounds were pre-draped, this limited the scope of the scenario design in terms of ability to introduce context and complexity. For example, one would usually check for neurovascular status of an injured limb prior to cleaning and draping it for suturing. We therefore developed more realistic “wearable” wound prostheses through collaboration with professionals from the film industry (Health Cuts Limited Ltd.) (Figure 4.3).

Figure 4.3: Hybrid simulation of wound laceration using wearable “prosthetic” sleeve to make wound appear to be part of the patient.

SP development

Professional actors with substantial simulated patient role experience were utilised for the study. Actors were trained to portray either the simple easy or difficult patients. Roles were standardised using semi-structured scripts and refined during pilot scenarios using video feedback (Appendix 4).

Clinical Facilities and Equipment

Instruments for performing all stages of the procedural skill such as gauze, syringes, needles, sutures and suture pack were identical to those used in real clinical practice and used in both PTT and PFHS scenarios. Patient's medical notes were also available. Simulations were conducted in a simulated clinical environment *Distributed Simulation,* which was previously validated to simulate an operating theatre (Kassab et al., 2011). The environment and equipment in this study were set up to represent a treatment room of an Emergency Department. In-built video recording technology was used for video ratings of performance.

4.2.5 Assessment of Performance

Participants were rated by 2 external independent senior EM trainees (SN and AN), who had no prior contact with the participants and were blinded to the level of training of the participant. Both raters were experienced and clinically competent at traumatic wound management, as well as trained in assessment methods. Ratings of the scenarios were conducted via the video recordings of simulations. These were carried out using existing validated instruments (Appendix 4) which were modified for the purpose of this study: Objective Structured Assessment of Technical Skills Task Specific Rating (OSATS-TSC; 10 items); Objective Structured Assessment of Technical Skills Global Rating Score (OSATS-GRS; 6 items); and the Direct Observation of Procedural Skills (DOPS) assessment tool (11 items) currently used in workplace based assessments (Martin et al., 1997; Norcini and Burch, 2007). The rationale for employing the OSATS rating tools were used for in-depth technical skills assessment, whilst the DOPS tool aims to assess global competence incorporating both technical and other non-technical skills and attributes such as professionalism and communication skills. Assessors were trained to use the rating tools specifically used in this study using video recordings from the pilot simulations. In addition, a process of calibration was performed whereby inconsistency of ratings and disagreements between the two assessors and questions relating to the application of the rating tools were discussed until an agreement was reached. Subsequent video ratings for the main study were conducted independently.

4.2.6 Data Analysis

Data analysis was conducted using IBM SPSS Statistics for Windows (Version 20.0. Armonk, NY: IBM Corp) software package. Internal consistency and inter-rater reliability of performance ratings were determined using Cronbach's alpha and intraclass correlation coefficient (ICC) respectively across OSATS-TSC, OSATS-GRS and DOPS. Mann-Whitney U test was used to compare performance ratings between experts and novices in both PFHS1 and PFHS2 scenarios. Wilcoxon matched-pairs, signed rank test was used to compare participants' performance ratings across PFHS1 and PFHS2 scenarios. A p value of < 0.05 was considered statistically significant.

4.3 Results

4.3.1 Demographics

Demographic information of the participants is presented in Table 5.1.

Table 4.1: Demographics of participants

Characteristics	Novices (n=10)	Experts (n=10)
Men (%)	20%	50%
Age (mean, SD)	26.4 years, 3.41	30.8 years, 4.39
Experience	<30 Traumatic wound management and closure	> 30 Traumatic wound management and closure
Experience with part task trainer simulation %	100%	100%
Experience with simulated (actor) patient %	100%	80%
Experience with hybrid simulation %	40%	50%
Postgraduate years, (mean, SD)	(1.2, 0.42)	(5.1, 2.02)

4.3.2 Performance Ratings

All rating forms, OSATS-TSC, OSATS-GRS and DOPS demonstrated high levels of internal consistency across both scenarios (ranging from 0.82 to 0.96). ICC was statistically significant (p<0.01) in all rating tools across both scenarios (Table 4.2).

Table 4.2: Internal Consistency Reliability and Intraclass Correlation Coefficients for the OSATS-TSC, OSATS-GRS, and DOPS across Scenarios and Raters

Scale	Number of items	Rater 1	Rater 2	ICC	*p*-value
PFHS1					
OSATS –TSC	10	.82	.81	.586	*p* =.003
OSATS - GRS	6	.95	.94	.819	*p* =.000
DOPS	11	.92	.95	.828	*p* =.000
PFHS2					
OSATS – TSC	10	.92	.90	.831	*p* =.000
OSATS – GRS	6	.95	.96	.857	*p* =.000
DOPS	11	.94	.95	.885	*p* =.000

4.3.3 Comparison between Expert and Novice Ratings

A Mann–Whitney U test was performed to test for differences in performance between the two groups (Table 5.3). The results of the test revealed that there was no difference on OSATS-TSC, OSATS-GRS and DOPS scores between experts and novices in PFHS1 scenario. Mean OSATS-TSC, OSATS-GRS and DOPS scores for experts were significantly higher than novices in PFHS2 scenario.

Table 4.3: Comparison of OSATS-TSC, OSATS-GRS, and DOPS scores between experts and novices across scenarios

	Novices		Experts			
	Median	**SD**	**Median**	**SD**	**U Value**	***p*-value**
PFHS1						
OSATS –TSC	8.25	2.19	9.50	0.85	66.50	*p* = .172
OSATS - GRS	3.10	0.71	3.65	0.80	73.00	*p* = .081
DOPS	3.30	0.63	3.87	0.90	71.00	*p* = .112
PFHS2						
OSATS –TSC	6.70	3.50	9.40	0.91	79.50	*p* = .023
OSATS - GRS	2.73	0.90	3.83	0.84	86.00	*p* = .006
DOPS	2.83	0.85	3.96	0.97	80.50	*p* = .021

4.3.4 Comparison between PFHS1 and PFHS2 Ratings

The Wilcoxon signed-rank test was performed to test for differences across scenarios within each group (Table 5.4). Novices' OSATS-TSC and DOPS scores were significantly lower in PFHS2 than in PFHS1. There was no difference in novices' OSATS-GRS scores between PFHS2 and PFHS1. There was no difference in experts' OSATS-TSC, OSATS-GRS and DOPS scores between HS2 and HS1.

Table 4.4: Comparison of OSATS-TSC, OSATS-GRS, and DOPS scores across PFHS1 and PFHS2 for experts and novices

	PFHS1			**PFHS2**			
	Median	**Range**	**Min-Max**	**Median**	**Range**	**Min-Max**	***p* value**
Experts							
OSATS – TSC	10.00	2.00	8.00-10.0	9.75	2.50	7.50-10.0	p =.317
OSATS - GRS	3.95	2.30	2.20-4.50	4.00	3.08	1.92 - 5.00	p = .161
DOPS	3.91	2.68	2.27-4.95	4.08	3.11	2.06 – 5.17	p =.059
Novices							
OSATS – TSC	9.25	5.00	4.00-10.0	8.75	10.00	0.00-10.0	p =.024
OSATS - GRS	3.15	2.20	1.80 – 4.00	2.96	2.83	1.00-3.83	p = .314
DOPS	3.32	1.91	2.23 - 4.14	2.89	2.67	1.50-4.17	p =.007

4.4 Discussion

In this study, I explored the use of PFHS for assessment of procedural skills for EM doctors. Specifically, I explored the potential of PFHS to recreate different levels of challenge by altering the clinical context in which a procedure, in this case the management and closure of a traumatic skin laceration, is performed. There were a number of key findings in this study. First, experts performed consistently in both simple (PFHS1) and complex (PFHS2), whereas novices' performance decreased significantly. This decrease in performance was detected using all three rating tools; OSATS-TSC, OSATS-GRS and DOPS.

When comparing novice and expert performance, there was no statistical difference between the two groups in the simple (PFHS1) scenario. Novices performed significantly worse than experts in the complex (PFHS2) scenario in terms of OSATS-TSC and DOPS ratings, i.e. the complex scenario discriminated between novices and experts. This supports the hypothesis that PFHS2 scenario was constructed to sufficient increase challenge such that a 3rd year EM registrar (expert group) should be able to complete the scenario competently, whereas more novice trainees might not.

Whilst the OSATS-TSC and DOPS ratings showed a significant difference in rates, the OSATS-GRS did not. The reason for this is uncertain, though a possible explanation is that the OSATS rating tools only measure technical performance and therefore do not take into account differences in non-technical performance between novices and experts, thereby limiting discriminatory power. When using simulations for assessment, not only do the simulations need to be valid, but also the rating tools must adequately reflect the purpose of the assessment. In this study, it is the assessment of an integrated performance of both technical and non-technical skills that is of concern. Therefore the DOPS tool, which is designed, assesses global performance, is arguably more suitable than the OSATS.

This study has a number of implications. Firstly, this is the first study to the best of my knowledge that investigates the use of PFHS to simulate different levels of clinical challenge of a procedure skill through systematic simulation design. The increase in level of challenge was objectively demonstrated where novices performed worse in PFS2 in comparison to PFHS1. Furthermore, it was demonstrated that the level of clinical challenge is an important consideration for procedural skills assessment – only the more challenging scenario (PFHS2) differentiated between experts and novices, demonstrating discriminatory ability. Interestingly, these findings reflect some of results of the second component of this empirical study which will be presented in Chapter 5, whereby in general novices found the complex scenario (PFHS2) to be more challenging whereas some of the experts did not.

The importance of recreating different levels of clinical challenge in simulation-based medical education has been highlighted previously (Mcgaghie et al., 2010). An assessment of insufficient level of difficulty may not be able to adequately assess a clinician's level of competence. However, there appears to be little in the current literature on how the level of clinical challenge of a simulation scenario or task should be set for the purposes of assessment. From my own experience in medical education and that of our research group, simulation-based assessments of procedural skills are often not sufficiently constructed to match a trainees expected level of competence and as such may be less useful and valid, particularly for high stakes exams. In addition to the simulation, the assessment tools used to rate performance must be equally robust. In this study, good inter-rater reliability across all rating tools was demonstrated.

There were a number of limitations to this study. Firstly, it was based on one clinical procedure, i.e. management and closure of traumatic skin laceration, so the findings are not necessarily generalisable to other clinical contexts. This is an exploratory study looking at the concept of introducing clinical challenge using hybrid simulation and future studies will need to look at a wider range of procedures. Secondly, the sample size is small with a lack of power study. This is a common issue with simulation studies, due to practical difficulties in recruiting participants. Finally, for reasons of research ethics, recruitment of participants was through voluntary participation, which could have introduced self-selection bias. However, both novice and expert groups were equally affected. Other potential biases such as differences in amount of prior exposure to simulation-based training are unlikely as the demographics data shows generally similar level of exposure between novice and expert groups.

4.5 Summary to Chapter

In this study, I described a systematic simulation design process, which included simulated patient development with use of wound prosthetics, equipment and facilities design, and scenario development based on cognitive task analysis to create PFHS scenarios of different levels of complexity for the purpose of assessment. This study demonstrated a controllable increase in the level of challenge of a procedural skill simulation – an important consideration in simulation-based assessment - through modifying clinical context using PFHS. The findings of this study are based on one clinical procedure at two levels of clinical complexity and is therefore not generalisable, and future work will need to expand on this study to investigate use of PFHS over a wider range of procedures and range of clinical challenge. However it is possible to see how this approach may be used to design simulations for other clinical skills and procedures at different levels of challenge to match a doctor’s level of training and expected competencies, with better alignment to training curriculums. When using these simulations for assessment it is also important to ensure that not only the level of challenge, but the rating tools as well are robust enough in order to provide a valid platform of assessing an individual's competence. In the next chapter, I will present the second component to this study, exploring the concept of contextualised simulation in PFHS from a different angle, specifically comparing the use of PTT-based simulations to PFHS.

CHAPTER 5 - Exploring Participants' Perceptions of PFHS to PTT-Based Simulations for Assessment of Clinical Skills

5.1 Introduction to Chapter

In the previous chapter, I investigated use of PFHS as an approach to introduce and modify clinical context to simulate clinical challenge. In this chapter, I will present the second component to my empirical study investigating the role of context in SBE which aims to answer key question 1B): How do clinicians perceive and compare simulation-based assessment with PFHS to PTT used in isolation? As mentioned in earlier, this component of the study was conducted concomitantly with the component presented in the previous chapter, for which an over all map of the study is presented in Figure 4.1. As with the first component of the study presented in Chapter 4, the simulated scenario of management of a traumatic skin laceration has been used as a platform for inquiry.

With respect to key question 1B, as discussed earlier in chapters 1 and 2, a key argument for using PFHS is to create a more authentic platform in relation to real clinical practice for the training and assessment of clinical skills by situating tasks around a (simulated) patient. There is at current some evidence that demonstrates user (trainee, assessee) satisfaction with PFHS for training and assessment across a number of different clinical scenarios. There also is some evidence that students and trainees find PFHS to be generally realistic. However, from my personal experience of SBE as a teacher and learner, students' behaviours may not necessarily always be authentic to real clinical practice. The relationship between simulation fidelity and participants' behaviour is not yet certain. Many simulations that are currently used in healthcare education contain elements that are clearly artificial, but may yet trigger authentic responses. Conversely, some that are seemingly higher fidelity may not necessarily do so. There is some interesting work by Dieckmann et al., (2007) who explore what they term as "reality" and "fiction" cues - elements of a simulation which are seen as realistic and unrealistic and may promote or inhibit user engagement. Interestingly, many elements of the simulations they studied were perceived differently from one clinician to another in terms of realism.

There is however, a paucity of literature that systematically evaluates how simulation participants truly perceive these simulations and the underlying mechanisms into their performance and behaviour within these simulations. In essence I am concerned with what is often termed in the domain of psychology as the "black-box". The black box represents underlying processes between intervention or stimulus and output or outcomes and is termed as such as it is sealed and inaccessible. In much of the current research in SBE, studies tend to focus on intervention and outcomes drawing what I believe are oversimplified causal links, e.g. use simulation X and outcome Y will be achieved. However, as discussed earlier in Chapter 3, simulation is a complex educational intervention that is multifactorial and multifaceted in nature. In order to truly understand how and why approaches to simulation work or fail, the underlying processes and mechanisms need to be investigated. This is particularly important for informing future simulation design.

5.2 Methodology and Methods

5.2.1 Study Design

The aims of this study are to elucidate clinicians' perceptions and experience of participating in, and being assessed with these two forms of simulation (PTT and PFHS). In order to answer this, as discussed in Chapter 3, a mixed-method approach combining qualitative (semi-structured interviews) and quantitative components (qualitative) was employed. The reasons for using mixed-methods are to provide differing yet complimentary approaches data that may allow for triangulation and cross-examination, the rationale and details of respective components of which are discussed.

5.2.2 Participation in Simulations

All participants completed three simulation scenarios, a PTT-based scenario and two PFHS scenarios based on management and suturing of a traumatic skin laceration in an Emergency Department. The scenarios were described in chapter 4 as follows:

The PTT-based simulation was based managing a patient with a minor traumatic skin laceration and consisted of commercially available wound suture pad (Limbs and Things,

Figure 5.1) and the necessary clinical equipment to complete the task. A written briefing was provided with instructions to complete the task (Appendix 2).

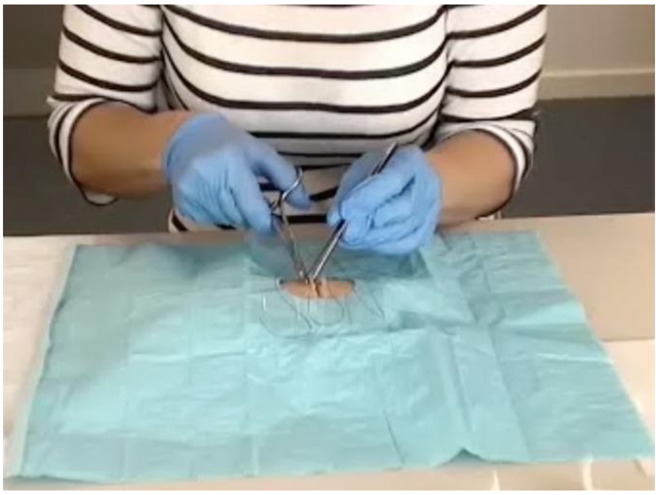


Figure 5.1 Participant suturing a wound on part-task wound suture pad

The two PFHS scenarios were designed to reflect the real-world practice of managing a patient with a minor traumatic skin laceration, that required suturing in an Emergency Department as described in Chapter 4 (Figure 4.3). These scenarios were designed to reflect two different levels of clinical challenge. The simulations were conducted in a previously validated simulated clinical environment to resemble an A&E cubicle (Kassab et al., 2011). All necessary equipment required to complete the scenario was available and presented in the same manner as in an A&E, i.e. in an equipment trolley. The equipment available was the same as that used in the PTT-based simulation. Patient's medical notes were also available. A wound prosthetic was placed on a simulated patient to resemble a traumatic laceration. The simulated patients were experienced patient actors who followed a semi-structured script (Appendix 3).

The participants were doctors across a range of clinical experience from foundation trainees to specialty trainees in Emergency Medicine (See Chapter 4). All participants had some previous experience of performing skin closure on real patients and/ or during some form of simulation exercise. Participants were recruited via e-mail and/or telephone communication.

5.2.3 Quantitative Component - Structured Questionnaires

For the quantitative component, its main purpose in this study is to obtaining measurable data to help identify general trends in participants' perceptions. From my personal perspective as an education and researcher, it can provide additional benefits in terms of providing accessible information for policy makers and other educators.

For this reason a structured questionnaire containing a series of Likert-type items was utilised. Seven items related to perceived realism in terms of the environment and behaviour. Eight items related to suitability of the simulation for demonstrating competence. The questionnaire was refined following a short pilot (3 participants). A questionnaire was administered immediately after each simulation scenario was completed.

5.2.3.1 Quantitative Data Analysis

Data analysis was conducted using IBM SPSS Statistics for Windows (Version 20.0. Armonk, NY: IBM Corp) software package. Statistical difference between participants' responses to PTT, PFHS1 and PFHS2 scenario was determined using a series of Wilcoxon Signed rank tests (related non-parametric comparison of medians). An alpha level of 0.05 was used to delineate statistical significance.

5.2.4 Qualitative Component - Semi Structured Interviews

For the qualitative component, the main purpose was to inquire in depth participants' perceptions of the simulations to elucidate possible underlying mechanisms and causal links to explain their experiences, as well as to compare and contrast with quantitative data. It also aimed to capture data that may not be readily obtained through the quantitative measures, in particular those that may be unexpected and outside my preconceived ideas as a researcher.

5.2.4.1 Rationale For Using the Interpretative Phenomenological Approach

Qualitative research is non-prescriptive with a multitude of approaches many of which have a degree of overlap (Maxwell, 2008). The approach employed should be inline with not only the aims of the study but intended outcomes. In this study, the aims are not necessarily to derive generalizable theories or test hypotheses, but to gain a personal insight into participants' experiences, as well as how it related to the existing understanding of the field. In this study, an Interpretative Phenomenological Analysis approach was used. Epistemologically, it is derived from three main areas of philosophy: phenomenology, hermeneutics, and idiography (Nollaig, 2011, chap. 3).

Phenomenology is concerned with an individual's perception and meaning of a their "lived experience" (Laverty, 2003). The philosophical origins of phenomenology stems from the works by Husserl in the early 20th century and with subsequent contributions by the likes of Heidegger and Merleau-Ponty which resulted in several schools of thought. For Husserl, his position of phenomenology is derived from the idea that knowledge and understanding of nature or essence of an experience is only possible when pre-conceived ideas and assumptions such as the nature of "existence" of an object are removed or bracketed. Research approaches in line with this are therefore "descriptive" and neither inductive or deductive(Laverty, 2003; Nollaig, 2011, chap. 3).

For Heidegger, phenomenology requires not only description, but also interpretation and hermeneutics – a key underpinning philosophy to the interpretative phenomenological approach. This is based on the notion that understanding of events, experiences or objects is mediated by the researchers' and participants' pre-existing knowledge and conceptions. However, interpretation may also be confounded by these pre-conceptions and therefore a careful balance between interpretation of the phenomenon of interest itself and how it relates to preconceptions needs to be achieved (Laverty, 2003; Smith et al., 2009, chap. 2).

The third major philosophical consideration underpinning the interpretative phenomenological analysis approach is idiography. The emphasis is to provide in depth understanding of an individual's unique "lived experience". The rationale and principles are related to that of narrative and case studies. As a research approach, it allows for gaining a unique perspective of each single individual's experience and how they related to the context of the persons in terms of their background, culture and ideologies (Nollaig, 2011, chap. 3; Smith et al., 2009, chap. 2).

Returning to the research question, by combining these principles together, this approach allows for both in-depth inquiry into individual's unique experiences, yet finding commonalities and differences in their shared perceptions of the two types of simulations of interest, PTT and PFHS.

5.2.4.2 Data collection

A semi-structured interview schedule was developed in line with the study aims. The areas of inquiry in the schedule was informed by both the relevant literature and also in part aligned to the questionnaire to compliment one another.

Interviews began with predominantly open questions to allow interviewees to freely express their thoughts, whilst closed questions were employed to clarify responses. No fixed questions were used, to allow flexibility of inquiry. A degree of modification was allowed for after each subsequent interview to incorporate themes and ideas that arose that were not previously considered.

The interview schedule was piloted with the pilot study in Chapter 4 with 3 participants.

Broad themes covered in the interviews after the pilot were as follows:

- Use of the different types of simulations for assessment of clinical skills
- Participants behaviour in the simulations
- Participant's motivation to perform in simulations.
- Perceived realism of the simulations
- Challenge and Difficulty

All participants were interviewed immediately after completing all 3 simulation scenarios (PTT, PFHS1 and PFHS2). Interviews were recorded using a digital audio recorder. Field notes were also obtained, which in addition to keeping thoughts of the interview were also used to capture data, which may provide more contextual information to the recorded audio such as hand gestures, laughter and so on.

The interviews were transcribed by professional transcribers (www.thetranscriptionagency.com), who specialise in medical transcriptions. The first 5 interviews were also transcribed by myself in order to validate transcriptions by the professional transcribers.

5.2.4.3 Qualitative Data analysis

The interviews were analysed inline with principles of Interpretative Phenomenological Approach discussed above. Each interview transcription was individually analysed both as single data sets in accordance with the principle of idiography prior to cross-analysis to identify commonalities.

Transcriptions were read in whole prior to coding. The coding process was iterative and modified as more interviews were conducted and analysed specifically for simplification and consolidation and developing emergent themes. Emergent themes were also generated with each new interview using a combined deductive (reflecting literature and pre-conceptions with comparison to quantitative data) and inductive process (to identify areas that are new to the researchers', i.e. my pre-conceptions). Two researchers (JKT and AG) initially worked collaboratively to generate early emergent themes through analysing 3 interview transcripts. Early in the analysis, interviews were also analysed in conjunction with the questionnaire findings for reasons described above.

As stated earlier, the Interpretative Phenomenological Analysis approach as with qualitative research in general is not prescriptive. Instead it provides a set of flexible guiding principles into inquiry and analysis dependant on both the aims of the research. In this study, the guiding principles set out by Smith et al (2009, chap. 5) were used to derive themes. This included abstraction (clustering of similar themes into a new common super-ordinate theme), subsumption (converting an emerging theme into a super-ordinate theme and bringing related themes under this), polarisation (consideration of differences and similarities) and contextualisation (drawing on contextual elements into the analysis).

5.2.4.4 Self-Reflexivity

Reflexivity is the process of acknowledging and reflecting of the potential impact of the researcher and his or her beliefs, attitudes and pre-conceptions on the study and how this may affect data collection and analysis (Yardley, 2000). This consideration is particularly pertinent in this study, given the approach to analysis used. I have therefore presented an account of my relevant personal background, to allow both the reader to gain an insight, and to encourage myself as the researcher to be mindful of the resulting possible influencing factors when conducting the study.

In terms of my position as a researcher, largely due to my background as a clinician and biomedical scientist, I have until recently mainly adopted the positivist stance in research, in that objectivity is the key to finding “truths” and that randomised-controlled trials are the gold standard for research. However more recently, largely due to completing a Masters degree in Clinical Education, I had the opportunity to gain some insight in the educational research, much of which is based on interpretivist paradigms. As a result, my position as a researcher has gradually changed from being purely positivist to one that also acknowledges and adopts the interpretivist ideologies.

In terms of my background as an educator, I have experience in both teaching and assessing in both formal and informal settings, e.g. lecture and classroom based versus ad-hoc ward-based. I am also aware of my personal experiences as a “consumer” of medical education both bad and good.

The area of medical education that I am most interested in is simulation and am a strong believer in its potential benefits. However, as simulation practitioner and having explored the literature quite extensively, I am critical of how it is used and researched. Specifically in relation to this thesis and the two types of simulation of interest, even though I am the designer of the PFHS simulations in this thesis, my position is that I believe neither is superior to the other. I do however have ideas of when one might be better suited than the other depending on the educational purpose as discussed in Chapter 2 and 3. These factors have influenced my motivation to conduct these studies and are reflected in my writing and analysis.

5.3 Results:

20 participants completed three simulation exercises (1 PTT-based and 2 PFHS) in total. A summary of the demographics is presented in table 5.1.

Table 5.1: Demographics of participants

Characteristics	**(N =20)**
Men (%)	35%
Age (mean, SD)	28.6, (4.44)
Experience with part task trainer simulation %	100%
Experience with simulated (actor) patient %	90%
Experience with hybrid simulation %	45%
Postgraduate years, (mean, SD)	3.15years (2.46)

5.3.1 Participants' Ratings of the Simulations

The participant's ratings for each question item are presented in table 5.2. Questions 1-7 relate to the perceived realism of the simulation and immersion. Questions 8-15 related to demonstration and assessment of competence. Comparison of the ratings between different simulations using Wilcoxon Signed Rank test is also presented in the table. * Denotes statistical significance.

Table 5.2: Participants Responses to Questionnaire Items Relating to Realism, Immersion and Suitability of PTT-Based simulations and PFHS for Assessment of Clinical Competence

	Median (Interquartile Range)			Comparison of ratings		
Question Item	**PTT**	**PFHS1**	**PFHS2**	**PTT vs PFHS1**	**PTT vs PFHS2**	**PFHS1 vs PFHS2**
Items relating to simulation realism and immersion						
Q1) This scenario in the simulation approximates closely to what happens in the workplace	3 (2.25-4)	4.5 (4-5)	5 (4.25-5)	P= 0.00	P= 0.00	P= 0.059
Q2) The wound is a realistic representation of a real wound	3 (2-3)	4 (4-4.75)	4 (4-5)	P= 0.00	P= 0.00	P= 0.659
Q3) The equipment used in the simulation is realistic	4 (4-5)	5 (4-5)	54-5)	P= 0.098	P= 0.08	P= 0.317
Q4) The environment is a realistic representation of the workplace	2 (2-3)	4 (4-5)	4(4-5)	P= 0.00	P= 0.00	P= 0.157
Q5) I was fully immersed in the simulation	3 (3-4)	4.5 (4-5)	4 (4-5)	P= 0.00	P= 0.001	P= 0.705
Q6) I performed as I do in the real workplace	3 (2-4)	4 (4-5)	4 (4-5)	P= 0.002	P= 0.00	P= 0.157
Q7) I behaved in the same way as I do in the real workplace	3 (2-4)	4.5 (4-5)	5 (4-5)	P= 0.001	P= 0.00	P= 0.257
Items relating to assessment of competence						
Q8) The simulation allows me to adequately demonstrate my technical skills	4 (3-4)	4 (4-5)	4 (4-5)	P= 0.002	P= 0.012	P= 0.48
Q9) The scenario allows me to adequately demonstrate my clinical knowledge for wound management and closure	3 (2-4)	4 (4-5)	4 (4-5)	P= 0.00	P= 0.00	P= 0.317
Q10) The simulation allows me to adequately demonstrate my aseptic technique	2.5 (2-4)	4 (3-4.75)	4 (4-4.75)	P= 0.001	P= 0.001	P= 0.248
Q11) The simulation allows me to adequately demonstrate my team working skills	1.5 (1-2)	4 (4-4)	4 (4-4.75)	P= 0.00	P= 0.00	P= 0.317
Q12) The simulation allows me to adequately demonstrate my professionalism	3 (2-3)	4 (4-5)	4.5 (4-5)	P= 0.00	P= 0.00	P= 0.083
Q13) The simulation allows me to adequately demonstrate my communication skills	2 (1-2	4 (4-5)	4 (4-5)	P= 0.00	P= 0.00	P=0.206
Q14) The simulation is able to highlight strengths and weaknesses in my workplace performance	3 (2-3.75)	4 (4-5)	4 (4-4.75)	P= 0.00	P= 0.00	P= 0.317
Q15) The simulation is an accurate judge of my overall competence in managing wounds	3 (2-4)	4 (4-4.75)	4 (4-5)	P= 0.00	P= 0.001	P= 0.257

Table 5.3: Superordinate themes, Subtheme and Sub-subthemes Derived from Analysis of Semi-structured Interviews

Superordinate Themes	Subthemes	Sub-Subthemes
Suitability and Acceptability for assessment of Clinical Competence.	Reflection of true competence	Holistic assessment
		Integrated performance Professional Context
	Should be Fit for Purpose	Objective of the assessment
		Developing component skills
		Feasibility, cost and time
	Negativity towards PTT	Lack of professional relevance
		PTT simulations in OSCE examinations
Authenticity of simulation experience	PFHS is realistic but	Variation in perceived realism
		Self-contradictory
		Human interaction and behaviour
	Authenticity and Behaviour	Behavior and performance
		Presence of a patient
	Limitations on authenticity	Knowledge of being in a simulation
		Limits built into simulation.
		Artificiality of role
	Relevance to Professional Practice	Meaningfulness Emotional Response
		Professional Identity
Challenging Simulation, Simulating Challenge?	Challenge of Simulation Difference between PTT and PFHS	Increased cognitive load
		Artificial setting of the PTT
	Difference in perceived challenge between PFHS1 and PFHS2	Contextual challenge
		Participant experience level

5.3.2 Findings From Semi-structured interviews

Analysis of the interview data yielded 3 super-ordinate themes and 9 subthemes, which were further divided into various sub-subthemes as presented in Table 5.3. The super-ordinate themes represent a broader insight of participants' shared experiences. Details of the findings with illustrative verbatim together with the level of expertise of the participant referenced are presented below. Sub-subthemes are presented in ***bold italics*** within body of the intervening commentary text. Although the aim of the analysis is focused on depth of inquiry rather than quantification, numerical data is provided where relevant to illustrate convergence or divergence in opinions and perceptions amongst participants. Level of expertise of the participant quoted (relative expert and novice according to the criteria used in Chapter 4) in the findings is also stated.

5.3.2.1 Superordinate Theme 1: Suitability and Acceptability for assessment of Clinical Competence.

This theme is related to participants perceives of use of the two types of simulations for assessing competence

Subtheme 1.1: Reflection of true competence

All participants felt that PFHS was an acceptable way of assessing their procedural skills competence.

> *"Its (PFHS) representing the situation, like representing what you do at work, because there's so much to actually being a doctor rather than just going in (and suturing)" (Participant 20, Novice)*

In addition, participants generally perceived the PFHS scenarios to be better for assessing a doctor's true competence in comparison to PTT. The underlying reason appears to be that the PFHS scenarios assessed competence more ***holistically*** in terms of requiring them produce an ***integrated performance*** of the required skills. PFHS also allowed participants to perform more similarly to how they would in a real clinical encounter.

> *"I think that the hybrid ones (PFHS) are a better demonstration of my clinical competence because of the prompts that, about the history, how the wound happened, whether you need to look at foreign bodies, likelihood of there being a deep penetrating injury, and contamination, age of the wound, all of those factors, which are just much more natural to ask about in, when you've got someone to ask compared with when you've just got a piece of rubber (PTT) and you have to try and list things off the top of your head" (Participant 07, Novice)*

> *"Not only do they allow you to demonstrate your procedural technical skills, but also other skills can be assessed which are part and parcel of that, such as professionalism and communication with patients, and team working and things. I think that they certainly feel very realistic as well.... In fact, it's fairly typical weekend evening patient in A&E." (Participant 05, Expert)*

In addition, by embedding the technical task (suturing a wound) within the wider ***professional context***, the PFHS scenarios provided certain nuances and potential distractions, not reflected in the PTT scenario, that a clinician may need to manage, thus better reflecting real clinical practice.

> *"… in a bench top situation, it's so controlled, you just suture. Whereas in a real life situation, the patient's moving, you have to talk to them. You're not going to be (just) focussing 100% on getting the suture placement right. There's going to be a lot of other things to distract you. And that's, it's probably a better test of what you would do in real life…" (Participant 13, Novice).*

Subtheme 1.2: Should be Fit for Purpose

Many participants discussed their opinions about the need for the types of simulations used to be fit for purpose.

A number of participants (60%) commented that part-task trainers have a key role in assessment of novices and for isolated skills. The reasons for this are two-fold. First there is recognition for the need for a platform for ***developing component skills*** as:

"You wouldn't expect someone who hasn't sutured before, who hasn't learned how to suture, to go straight into a hybrid simulation test" (Participant 06, Expert)

The type and design of simulation used should also reflect the ***objective of the assessment*** or assessment criteria, and not just be made as authentic or realistic as possible for the sake of it.

"I think if you're just assessing suturing skills you can do it on that model… of if you just want to teach people to suture correctly, there's nothing wrong with using those and they're very useful actually." (Participant 19, Expert)

"… if you were literally just assessing someone's ability to do a communication skills or to suture, you should probably test those in isolation." (Participant 15, Novice)

In addition to discussing the appropriate type of simulation for its desired purpose, a number of participants (30%) also raised the issue of ***feasibility*** of using PFHS routinely for assessments in terms of ***cost and time***, as well as some scepticism towards the practicalities.

"… but I'm just wondering, the difficulty obviously is the time it takes to set it up, and the cost and things, isn't it, presumably. How much do these prosthetics cost?" (Participant 05, Expert)

"But then I wonder how easy it is to translate to exam situations because of the cost. Because I know from a university point of view they won't do that just because it costs so much to do". (Participant 09, Novice)

Subtheme 1.3: Negativity towards PTT

Interestingly, even though participants acknowledge the usefulness of PTT simulations as seen in the earlier theme, there appeared to be a degree of negativity towards the PTT simulations by a number of participants (40%), as one commented:

> *"The bench top one (PTT) , it almost seemed a bit, almost farcical" (Participant 13, Novice)*

Part of the reason for this negativity may be the perceived ***lack of relevance*** to professional practice by some participants

> *"(its) just like playing like a game (Participant 17, Expert)*

> *"I just felt like it was a practical procedure and I was just doing it for practices sake" (Participant 18, Novice)*

> *"it feels a bit abstract on the part task" (Participant 12, Novice)*

There also seems to be quite a strong influence from participants' past experience of being assessed with PTTs. Specifically, they seem to be triggered from memories of negative experiences of ***PTT simulations in OSCE examinations***. The negative experiences not just the PTT simulation, but how the OSCEs themselves were conducted. For the PTT simulation, participants stated how they found it difficult and award to interact with a PTT.

> "I've always hated OSCEs from that point of view of having to talk to, oh I don't know, a fake wound or a fake bum for a PR exam or something like that." (Participant 09, Novice)

At the same time, some participant reflected on how their experience with PTT simulations in OSCEs were more akin to a tick-boxing exercise.

> *"I am trying to remember my OSCE skills more rather than actually just getting on and you know (performing)"(Participant 11, Novice)*

> *"When you do exams now, you learn to go almost in automatic mode, so, I'm going to tick these boxes..." (Participant 04, Expert)*

Summary of Super-Ordinate Theme 1

In summary, participants acknowledged the usefulness of both types of simulation, though PFHS is perceived to be better reflection of true competence. Central to this is that it allows a clinician to demonstrate their skills in a manner similar to in their professional practice, i.e. in a holistic integrated fashion. Issues were however raised about cost and feasibility of PFHS in routine practice. With respect to PTTs, although perceived to be useful, there was a degree of negativity towards its use. Part of this appears to be related to this type of simulation itself where some participants found a lack of relevance to professional practice, resonating with issues with decontextualised tasked discussed in Chapter 1.

The negativity also seemed to be significantly influenced by past negative experiences of being assessed with PTT in the context of an OSCE exam. This raises questions about wider implications of assessment in general in terms of unexpected effects will be discussed in more detail later.

5.3.2.2 Superordinate Theme 2: Authenticity of the Simulation Experience

This super-ordinate theme relates to participants perceived realism of the two types of simulation in terms of reasons for, and potential effects of. All participants perceived the PFHS scenario to be highly realistic and authentic to real clinical practice. In comparison, all participants perceived the PTT scenario to be unrealistic.

> *"…the (PFHS) simulation is like what it would be like in a real A&E situation…" (Participant 13, Expert)*

> *"…It feels a bit abstract on the part-task … it doesn't feel like treating a patient…"(Participant 12, Novice)*

Subtheme 2.1: PFHS is really realistic, but…

Although participants generally stated that they thought the PFHS scenarios were authentic to clinical practice, most perceived some parts of the simulation to be unrealistic. Interestingly many participants were ***self-contradictory*** *(see third adjoining quotation)*, on the one hand initially saying the PFHS scenario was true to life, but then on further questioning commenting on the unrealistic elements. Some also seem seemed have a dismissive attitude towards the unrealistic elements. This could reflect their expectations of how realistic a simulation can be, perhaps from past experiences. It may also reflect their ability to “suspend disbelief” and buy into the simulations. They were also aware that they were getting into “role” and acknowledged the presence of “actors”.

> *“I think was pretty realistic actually. Clearly I’ve walked into an “Igloo” (DS environment) and I’m inside some kind of weird bubble thing, so that takes away from a bit of it and it’s not the surrounding noise of the A&E and things like that, but I thought it was, yeah, pretty realistic.” (Participant 01, Novice)*

> *“I thought the actors were very good and the wounds looked very realistic and, yeah, it was quite strange that you do immerse yourself quite quickly in the situation and try and get on.” (Participant 08)*

> *“…it was quite realistic having, the wounds themselves are fairly well positioned, and look like they are almost part of the limb, and at times I actually did forget that it wasn’t the patient’s genuine skin”… (Earlier in interview)…and just minor things like when injecting the Lidocaine (local anaesthetic) because of the silicone it’s worked straight back out the hole, which made me laugh, and wasn’t terrible realistic...”(Later in the same interview, Participant 07)*

Some items were more consistently seen as unrealistic, though in general there was marked ***variation in perceived realism of the different components*** of the simulation. For example, 30% of participants specifically commented that they found the prosthetic wound to be realistic whilst 20% stated that they found the wound to be unrealistic.

> *“The wound itself looks realistic… the skin, the sort of give and feel of the suture feels more realistic (Participant 05, Expert)*

"Parts (of the simulation) are slightly unrealistic…. the texture and the feeling of the skin is going to be hard to approximate, and that didn't feel like normal skin to me." (Participant 06, Expert)

Other items that were perceived to be unrealistic ranged from the background environment and the presence of a simulated nurse.

"Definitely louder (in real life), and I think more, because you've got, obviously got hospital machine noises but more noise, more patient noise, people crying out, that kind of thing, because I've never heard an A&E that quiet really."(Participant 09, Novice)

I think in reality you tend to get interrupted a lot, particularly as a senior registrar. And people tend to come and find you every two minutes and ask you questions, which make doing a procedure like that even more difficult. But apart from that, everything else was much more realistic (Novice).

" Most of the times you don't have a nurse to help so you have to do everything yourself" (Participant 04, Expert)

Presence of a "patient" was however consistently seen as the key aspect that contributed to authenticity of the PFHS. In addition to providing cues and prompts that appear to encourage participants to put themselves in the mind-set of dealing with a patient, ***human interaction*** seems to be key.

"The realism, I think it's the human contact, it is the most important thing" (Participant 17, Novice).

"The fact the patient's talking to you all the time, the patient's moving... you forget that, actually, it's all staged and it's quite, all quite convincing" (Participant 03, Expert)

"It felt a lot more natural, and there were sort of the prompts of having the patient and the history in front of you made you think about elements of the wound management that were important." (Participant 07, Novice)

Subtheme 2.2 Authenticity and behaviour

Most participants (90%) stated that they were able to ***behave and perform*** as they do in real clinical practice in the PFHS scenarios, but they were not able to in the PTT scenario. As with perceived realism, the key to authentic behaviour appears to be the context of the clinical scenario and the ***presence of a patient*** within the PFHS simulations. Conversely, in the PTT simulations, the absence of a "patient" seems to negatively impact on participants' ability to perform authentically.

> *"I think I truly performed (in the PFHS scenarios) as I would in real life. I interacted with the patients in the same way. I didn't think this is an actor. I was in it the whole time" (Participant 08, Expert)*

> *"… you have got no emotional contact or sense of connection with a bit of plastic arm (PTT) on a bench top. Whereas in the scenario you've got the sort of stress of human interaction and the sense that this is actually somebody's arm and the extra pressure of wanting to make this look OK as this is someone's leg… I want to make this look good." (Participant 16, Expert)*

> *"When I was doing the part-task one… I wasn't behaving as I would in a clinical environment and I wasn't interacting with a piece of sponge (suture pad) in the way that I interact with a patient…" (Participant 07, Novice)*

Whilst most participants commented that they could generally behave in an authentic manner in the PFHS scenarios, some participants stated that the extent to which they could exhibit realistic behaviour was limited. In particular, the ***knowledge that they were in a simulation***, and that they were being observed seemed to be a major contributing factor.

> *"We interact with them (simulated patients) in a simulatory way, I won't be so touchy feely as you know you are being observed and its' not the real thing. (Participant 04, Expert)*

Some participants commented on inauthenticity that may be brought about by the some inherent ***limits built into simulation***. For example in as suggested by the following participant, although the SP in PFHS2 scenario was designed be challenging and abusive, he knew that there would be limits to what the SP would actual do for safety and ethical reasons, unlike in the real clinical workplace.

> *"I think realistically, I know that it's not somebody that's actually going to punch me, which is perhaps a limitation of the scenario." (Participant 10, Expert)*

Authenticity was also limited by the study design. In this study, relative novices were put through all the simulations including the more challenging scenario, PFHS2. The participants were also briefed to "step-up" into an ***artificial role*** of a more senior clinician and tackle the scenario in that role. However in real practice, the relative novices may not actually deal with such a patient themselves or at the very least may request senior input.

> *"And also that at my level obviously, I'd probably have a senior person who I could refer to about little questions" (Participant 11)*

Subtheme 2.3 Relevance to Professional Practice

Most participants (95%) were highly motivated to perform well in the PFHS scenario, whereas, participants were less inclined to perform well in the PTT scenario. The reasons for this difference in motivation appear to be related to the ***"meaningfulness"*** of the simulation as well as the drive to perform well on a "patient". The presence of a "patient" seemed to be able to elicit ***"emotions"*** relating to real clinical practice in some participants. Likewise, the absence of this emotional element was readily identified by some. Finally, being able to relate to their ***professional identity***, i.e. feel like a doctor, seemed to play a positive role in motivation.

> *"I was more motivated to do a nicer job and to make sure every stitch was good on the hybrid than on the bench top. Because it's almost like it belongs to a patient so it's more, it feels more important than just a bit of meat on a bench top." (Participant 14, Expert)*

"…if you're just suturing an arm, a bench top thing, you have got no emotional contact or sense of connection with a bit of plastic arm on a bench top" (Participant 17, Novice)

…if you're just suturing a bench top thing (PTT), you have got no emotional contact or sense of connection with a bit of plastic arm on a bench top. Whereas in the (PFHS) scenario you've got the sort of stress of human interaction and the sense that this is actually somebody's arm and the extra pressure of wanting to (do a good job).." (Participant 16, Expert)

"The bench top one (PTT), I'm not very motivated to do it, it's a bit of a chore. It reminds me of my clinical skills when I was a medical student…. as a doctor often, the reason you've gone into it is often because you enjoy talking to people and sorting problems out…. (Participant 02, Novice)

"…its (PFHS) representing the situation, like representing what you do at work. because there's so much to actually being a doctor" (Participant 20, Novice)

"The only thing is like, with some of the other models and things, I'm not quite sure, are you allowed to, how much you're allowed to inject" (Participant 03, Expert)

Summary to Super-Ordinate Theme 2

In summary, participants found the PFHS to be much more realistic than the PTT scenarios, and to be realistic in relation to real clinical practice. However, perceived realism of the different components is variable. Although participants seemed to be instinctively aware of the various nuances and artificialities of the PFHS, they were still able to get into role. This may be in part related to the presence of a human "patient", but may also be related to participants' expectations of what the simulations may be able to achieve.

Factors encouraging participants to exhibit authentic behaviour include not only the physical and contextual realism of the simulation, but the simulations' ability to elicit certain emotions and allow participants to relate to their professional identity. There are inherent limitations to the degree of authenticity that can be achieved, including an observer effect, and participants' awareness of being in a simulation.

5.3.2.3 Superordinate Theme: Challenging Simulation, Simulating Challenge?

This theme related to participants' perceived challenge between the various simulation exercises.

Subtheme 3.1: Difference in perceived challenge between PTT and PFHS scenarios

Doctors were divided in their opinion with respect to difference in challenge of the PTT and PFHS scenarios. Half of the participants commented that the PFHS scenarios were more challenging. The reason for this was the ***increased cognitive load*** required to simultaneously integrate a range of skills.

> *"I did find the hybrid one's more challenging… you've got more things to think about and you're talking to the patient whilst you're also trying to suture their arm up. So you're multitasking…"(Participant 02, Novice)*

The other half of the participants felt the PFHS scenario was less challenging than the PTT. The reasons for this appear to be that the PFHS scenario allowed participants to perform in a more naturalistic manner. The perceived relative ***artificial setting of the PTT*** simulation also appeared to make the task more challenging, potentially raising wider issues about use of decontextualised simulations.

> *"I think the bench top (PTT) is more challenging, because when you just talk to a patient, it kind of comes a lot more naturally (in the hybrid scenario)" (Participant 20, Expert)*

> *"I actually found the bench top more challenging, which is probably different for most people. Because it's just because you're by yourself, you're in this like very, very artificial environment." (Participant 08, Expert)*

Subtheme 3.2: Difference in perceived challenge between PFHS1 and PFHS2

Most participants (70%) generally perceived the PFHS2 scenario to be more challenging than the PFHS1 scenario. Reasons for this are predominately related to the ***context*** in which the procedure is conducted.

> *"In the second scenario (PFHS2), the gentleman was much more challenging. Clinically, he wouldn't keep still, he kept asking questions and interrupting, and was also a bit derogatory ... rude and confrontational, so that made it quite challenging." (Participant 07, Novice)*

However, a number of participants (25%) did not perceive any noticeable difference in challenge between the two scenarios. Further analysis revealed that this perception was mainly amongst more ***experienced participants***, i.e. senior EM registrars.

> *"Generally the people that I suture are either intoxicated and have fallen over, so I'm used to dealing with patients like him (PFHS2 scenario)." (Participant 06, Expert)*

Summary to Super-ordinate Theme 3

In summary challenges from simulation arise from a number of factors. Difference in participants' perceived challenge existed between both the types of simulation (PTT versus PFHS) and the complexity of the simulation (PFHS1 versus PFHS2). Whilst some participants found PFHS more challenging than PTT due to increased cognitive load, others found the PTT to be more challenging due to its relatively artificial nature. Difference in perceived challenge between the simple (PFHS1) and difficult (PFHS2) scenarios appeared to be related to participant's level of expertise, i.e. experts did not always find PFHS2 more challenging than PFHS1.

5.4 Discussion

With the increasing adoption of SBE, there is also a recognised need to increase our understanding of the underlying mechanism of how simulations work to encourage rational development and usage (Issenberg et al., 2011). Earlier, I presented the rationale for the use of PFHS as a means to creating a more holistic assessment of clinical competence. In this component of the study, I used a mixed-method approach to explore clinicians' experience of PFHS and PTT simulations for assessment of competence in detail.

There were a number of key findings in this study. First, the participants' ratings for questionnaire items relating to perceived realism of the simulations were generally high for the PFHS simulations (median ratings ranged from 4 to 5 for both PFHS1 and 2) and moderate to low for the PTT simulation (median ratings ranged from 2 to 3). Ratings relating to realism were significantly higher for most (6 of 7) questionnaire items for PFHS1 and all items for PFHS2 when compared to ratings for the PTT simulation.

Participants' ratings of questionnaire items relating to suitability for use in assessment of clinical competence was also high for the PFHS simulations (median ratings ranged from 4 to 4 for PFHS1 and 4 to 4.5 for PFHS2) and variable for the PTT simulation (ranging from 1.5 to 4). Further analysis, shows that only item Q8 which relates to technical skills is rated high (4 of 5) for the PTT simulation, whereby the remaining items are rated moderate to low (1.5 to 3). Ratings for items related to assessment and demonstration of clinical competence were also statistically significantly higher for all items for both PFHS simulations in comparison to the PTT simulation. Ratings for PFHS1 and PFHS2 simulations were similar and in particular, no significant difference was found between ratings for all question items relating to realism and those relating to assessment of competence between PFHS1 and PFHS2 simulations.

Findings from the semi-structured interviews relating to suitability and realism of each type of simulation generally correlated with findings of the questionnaire responses. In general, the participants perceived the PFHS scenarios to be highly realistic in relation to real clinical practice and allowed participants to demonstrate they clinical competence holistically. Conversely, the PTT-based simulation was generally regarded as artificial, although clinicians' also acknowledged their importance with respect to focused training and assessment of component tasks, particularly for relative novices.

Whilst both the questionnaire ratings and the interview responses regarding the authenticity and realism of the PFHS were high, more in-depth questioning revealed that realism of the different individual elements of the simulations such as the physical aspects and the scenario itself were perceived variably by participants. There may be a number of reasons for this. Firstly, whilst we attempted to provide an authentic clinical experience, it was not possible or practical to reproduce all the elements accurately. For instance, with respect to the physical elements, the wound did not bleed and was clearly "worn" by the patient.

Likewise, with respect to the scenarios, for the purpose of the study, all participants including novices were requested to assume the role of a more experience clinician and complete the PFHS2, i.e. the more challenging simulation. However, if they encountered this sort of scenario in their actual clinical practice, realistically they may choose to refer this case to someone more senior or ask for senior input, as one participant stated, *"I'm just not as experienced with suturing wounds, so I'd actually probably ask for help quite a lot more, or at least a level of supervision initially"* (Participant 07). As a result a degree of artificiality was introduced. However, this could be both and beneficial and detrimental depending on the purpose of the assessment. If the aim of the assessment aims to capture authentic performance at their an assessees' expected level of training such as for a high stakes exam, then it is arguably unfair to put them through more challenging scenarios. However, if the aim of the simulation is to put a relative novice in a challenging situation that they would than they would usually be asked to, in order to assess how they might cope, or to help them see how they can manage such situations if encountered in future practice, then it may have some educational value. As discussed in Chapter 1, simulations can be designed to be skewed representation of reality, sacrificing realism for the desired education outcome.

Another reason why there is variation in perceived realism is that the participants may differ in their ability to suspend disbelief. It is also generally known that peoples' interpretations and perceptions of representations can vary greatly and are highly subjective; some may look at a picture and find it to be realistic whilst others do not, as suggested by a long-established literature on perception bias (Hastorf and Cantril, 1954; Leuba and Lucas, 1945). The findings of this study are also in keeping with those of Dieckmann et al., (2007) as discussed earlier who also found variation in the perceived realism of different simulation components. Interestingly, during the interviews, many participants provided contradictory statements of their perceived realism of the simulations, on the one hand claiming the PFHS to be realistic, but on the other hand pointing out what were clearly quite unrealistic elements. At the same time, many were quite readily dismissive of the importance of the unrealistic elements giving the impression that it perhaps did not matter too much. The reasons for this are not certain, though there are several potential explanations. First, it may be related to the participants' expectations. All participants have had some experience of the comparatively less realistic PTT simulations prior to the study and also during the study potentially influencing their ideas of how realistic a simulation should or can be. Second, this phenomenon could be related to participants' ability to suspend disbelief in order to "get on" with the simulation.

Thirdly, despite some artificiality, when the elements of the simulation are combined together, there may have been an overall Gestalt effect resulting in overall perception of high realism, where, "the whole is greater than the sum of its parts" (Steinberg et al. 2007). In the PFHS simulations it may be that the overall representation of the scenario presented enough cues such as the wound, environment, patient and equipment that make it feel like an authentic clinical experience.

Finally, it may be that participants place different values on the different components of the simulation in terms of their realism. Where as the realism of certain aspects such as the background may not have been too important, others such simulating realistic patient interaction seemed to be a significant contributing factor to immersion in this study. This raises some interesting questions about simulation design, particularly with respect to fidelity requirements.

The two key aspects that were consistently perceived to be central to the realism and authenticity for the PFHS, was patient interaction and the clinical scenario in which the wound management procedure was embedded. These also appeared to contribute to the participant's motivation, behaviour and attitude towards performing in the simulation. A number of participants commented that in the PFHS scenario they performed almost as if they were performing on a real patient. Specifically, they felt more inclined to perform the procedure well due to the presence of a "patient". This is unsurprising as patient interaction is central to a doctor's daily work, and related to their "professional identity" particularly for procedures such as the one in this study, i.e. closure of a skin laceration under local anaesthetic. As a number of participants stated, "*it's like doing your normal job… you just get on with it*". In essence, the presence of a patient appears to make the simulation more meaningful.

On the contrary, many participants stated their lack of motivation and engagement with the PTT simulation and lacking relevance to professional practice. Despite participants acknowledging the usefulness of PTT simulations, many seemed to exhibit a degree of negativity towards them. Several participants commented on how the PTT simulation reminded them of their experience with OSCE exams which emphasised on completing necessary steps and “ticking boxes”. A quick search on any mainstream Internet search engines on OSCE will produce a host of webpages with negative comments of OSCEs from medical students and doctors. These opinions appear to relate not only to the simulations used in the OSCE, but also the assessment format and assessment criteria, which usually consist of checklists (Hodges et al., 1999). This raises some important issues. First this prior negative experience may have implications on study findings resulting in participants devaluing the usefulness of PTTs. Perhaps more importantly though, is the wider question of the unintended effects the assessments we carry out on our trainees. Few people may claim to genuinely like being assessed, though ideally they should appreciate the relevance of the assessments if deserved (Frymier and Shulman, 1995). Reflecting on my experience as an examiner and drawing on the principles of the authentic education paradigm discussed in Chapter 1, there is perhaps a need for greater efforts to ensure performance based assessments in general such as these have value beyond assessment and designed to better reflect professional practice.

Returning to the issue of realism, despite the PFHS being perceived as highly realistic there were a number of barriers to authenticity beyond the artificialities of the design of the simulation itself, such as the simulated environment or the prosthetic wound, which are inherent to simulation-based assessments in general. For example, one participant stated that she behaved in a “simulatory way”. Another participant stated that they adopted a more “softly” approach towards the aggressive patient in the PFHS2 simulation, whereas in reality, they would be sterner, to the point and perhaps even rude. Indeed, one of my observations whilst conducting the PFHS scenarios is that a number of participants shouted out loud their intended actions and thoughts, such as *“I will NOW WASH my HANDS”*, often at a camera. This behaviour seems to be similar to that of when people take driving tests, where they exaggerate actions such as checking the mirrors to ensure their assessor’s know they are performing the required manoeuvres in order to pass. In the setting of an assessment, assessees may be aware of the need to "demonstrate" their competence and therefore it is perhaps unsurprising that they would shout out their thoughts and actions. A look at some of the commercially available examination preparation books for OSCEs and similar types of performance-based assessments do advice students to speak out key words and phrases. From my experience, this behaviour of exaggerating actions to demonstrate ability is not limited to assessments with video performance, but exists in general when students are aware of being assessed, including in during workplace based assessments (Batchelder and McCarthy, 2013; Schuwirth and van der Vleuten, 2011). This issue of this Hawthorn effect in simulation-based assessments is by no means a criticism as "shouting out" may be the most suitable way for an assessee to convey underlying cognitive processes, but it is something that assessors and simulation designers should be aware of.

From the perspective of simulation design, there is some literature documenting approaches to minimising the Hawthorn effect in order to capture authentic behaviour and performance in simulation-based assessments by use of incognito SPs. In this type of simulation, SPs attend a clinic where the clinician is unaware that they are not real patients (Gorter et al., 2002). Of course, there are limitations to this as well, in that it does not allow capture performance of clinical skills that may be invasive or intimate.

On exploring how doctors compared PTT to PFHS scenarios in terms of challenge, participants were divided as to which they found more challenging. A number of participants suggested that PFHS scenarios were more challenging, due to an increase in cognitive demand due to the need to simultaneously perform several skills. On the other hand, many participants felt that the PTT were more challenging due to the unnatural setting in which the task is presented. There were indeed a few incidences I observed during the study where participants seemed to under perform on the PTT simulation. For example, one of the more senior participant's who was clearly usually clinically competent, omitted several key steps such as failing to clean the PTT wound and administer local anaesthetic prior to closure of the wound despite having a clear briefing instructions and the necessary equipment provided in front of her. Halfway through the simulation, the participant mumbled a curse in realisation, verbally apologised to the part-task trainer and requested to restart the simulation. During the post simulation interview, the participant stated that the reason might have been due to the artificial setting that made her forget to do things that come naturally to her.

Another incident I observed was a participant who was again quite competent, but kept repositioning the PTT to make it easier for him to suture the simulated wound. However in doing so, his aseptic technique was compromised. This raises important issues in terms of the extent to which PTT scenarios can allow doctors to demonstrate their true abilities. In essence, the level of challenge of the simulation experienced by doctors may at least in part be related to the format of the assessment rather than the clinical difficulty of the scenario. This is also an issue raised in the current literature on authentic assessment (Montgomery, 2002).

Looking at participants' perceived difference in the level of challenge in the two PFHS scenarios, the majority of participants generally found PFHS2 to be more challenging. The reason for the increase in challenge was the clinical context in which the procedure skill had to be performed, i.e. on a drunk uncooperative patient. However, a number of more experienced doctors stated that they did not find this context more challenging as they were experienced in this sort of encounter in their usual practice. This supports the argument that clinical context is an important consideration in simulation-based assessments, especially when ensuring the assessments are of a suitable level of challenge for the trainee. This was also in keeping with the findings presented earlier in Chapter 4.

In exploring participants' perceptions of the use of PTT and PFHS simulations for assessment, doctors agreed that the PFHS scenarios were better at eliciting true competence in comparison to PTT. The reason for this seems to be that PFHS assesses competence holistically taking into account a range of professional skills. Whilst participants generally believed the PFHS was a good approach for assessing clinical competence, there were issues raised as to the feasibility and costs, two important considerations in the utility of an assessment. Further studies will need to investigate these issues (van de Vleuten, 1996).

One of the criticisms of current “in-vitro” simulation-based assessment is that it does not necessarily predict a doctor’s real world “in-vivo” performance in practice as shown in previous studies. A number of factors may also influence a doctor’s performance in simulation. Current assessment theory based on Miller’s pyramid described in Chapter 1 (Figure 1.1) clearly delineates the upper hierarchies of “Does” which relates to in-vivo performance and “Shows how”, i.e. simulation-based in-vitro assessments. The reasons suggested include differences in context; one being in the laboratory and the other in the real clinical workplace (Rethans et al., 2002). However, findings from this study suggest that, by introducing more context, it may be possible for doctors to put themselves in an appropriate frame of mind so that their behaviour and performance better approximates what happens in real life. In also allows for the assessment of doctors’ competence in an integrated fashion, more closely resembling real world performance. At the same time, as mentioned earlier, it may not always be possible to capture “authentic” behaviour through direct observation of performance in the workplace due to a Hawthorne effect. I therefore argue that the boundaries between the upper two hierarchies of Miller’s pyramid of “Shows How” and “Does” are perhaps more blurred than is usually considered in the literature, and should be taken into consideration when determining the strategies for assessment of competence.

There are a number of limitations to this study. First, for practical reasons, as mentioned in chapter 4, the number of participants was low. In addition, the study was based on a single clinical procedure and therefore, findings may not necessarily be generalisable to other clinical procedures. The reason for this was primarily practical in terms of resource and time availability. Secondly, the participants were recruited voluntarily, thereby potentially causing self-selection bias. However, participants were recruited from a range of hospitals and at different training grades. In addition, participants were not informed of the types of simulation including the clinical scenario they would participate in prior to the study. Finally whilst this study aimed to evaluate PTT and PFHS-based assessments, the study itself was not conducted in the setting of an “official” high-stakes assessment. This may have effect on participants’ performance, for instance, they may be less prepared for an assessment, or have less incentive to perform well.

5.5 Summary to Chapter

In summary, in this chapter I presented the findings of the second component of a larger study, exploring context in simulation-based assessment with PFHS. Clinicians generally perceived the PFHS simulations in the study to be authentic to real clinical practice and a suitable method for holistic assessment of competence. They also perceived PFHS to be better means of assessing overall clinical competence in comparison to PTT-based simulations. In addition to just being providing a more realistic representation to real clinical practice, there may be additional benefits from the perspective of alignment with professional identity and meaningfulness. However, clinicians' also acknowledged the importance of PTT-based simulations, particularly with respect to focused training and assessment of component skills.

Various elements of the PFHS simulation were perceived more realistic or artificial by the different clinicians, though patient interaction and the context of the clinical scenario were consistently perceived by all participants to be central to authenticity. Patient interaction and the clinical scenario itself were also key to encouraging authentic behaviour. However, limitations in design of simulation-based assessment, and awareness of being assessed appeared to diminish the extent to which participants exhibit their true behaviour in clinical practice. With respect to PTT simulations, certain artificial aspects may detract participants from performing authentically. An understanding of these issues is necessary when designing PFHS based assessments, as well as drawing inferences of clinicians' performance within them.

CHAPTER 6 - Comparing Use of PFHS to Patients for training and Assessment of Clinical Skills

6.1 Introduction to Chapter

In this chapter, I will present the second main empirical study of my thesis. Earlier in chapter 1, I presented some of the key arguments supporting the use of simulation in general and more specifically PFHS for training and assessment. Much of current simulation training focuses on component tasks. Yet clinical competence requires the seamless integration of a multitude of competencies including skills from both technical and non-technical domains. In this respect, PFHS which integrates both non-technical and technical skills may offer some educational benefits. However, despite some compelling theoretical arguments, there remains limited empirical evidence to support its use as demonstrated in the systematic review presented in Chapter 2. There remains little direct evidence comparing training and assessment of clinical skills with PFHS to real patients. From the perspective of learning, there is a need to investigate if, and to what extent do skills learnt in simulation transfer to real clinical practice. At the same time, there is a need to ensure that it does not promote untoward learning effects. From the perspective of assessment, there is a need to know the extent to which inferences drawn from simulation-based assessments relate to real clinical performance. A further area that also requires additional research also discussed in Chapter 2 and 5, is that there is at current little in the literature with respect to what happens in the classroom in terms of the process of learning and underlying influencing factors.

The aim of the study presented in this chapter is to answer the following key research questions:

2A What is the process of learning with PFHS and how does this compare to learning with real patients

2B How do students training with PFHS compare to those training with real patients

2C How does assessment of clinical skills with PFHS compare to assessment with real patients

6.1.1 Rationale for Investigating PFHS Using the Training and Assessment of Cardiovascular Examination Skills

For the purpose of this inquiry, I chose to use the teaching and assessment of cardiovascular examination skills training of medical students as a platform to investigate PFHS. There are several reasons for this. First, from a clinical perspective, it is a commonly performed examination with demonstrated clinical utility (Roldan et al., 1996). When performed competently, cardiovascular examination can aid diagnosis and elucidate important pathology. It is a vital skill that all medical graduates are required to perform to a reasonable level of competence. Yet, there is a well-recognised decline in general level of competence of cardiovascular skills amongst doctors internationally and across different specialities (Alam et al., 2010).

The underlying reasons for this are multifactorial and reasons suggested in the literature include the increasing reliance on diagnostic technologies such as echocardiograms and reduction in requirements by training bodies to ensure competency (Alam et al., 2010). It has also been suggested that there may be a reduction in opportunity for developing competency in the workplace based due to inconsistent and haphazard training (Asghar et al., 2010).

The implications of general competency levels of cardiovascular examination skills can have on patient care range from misdiagnosis (both over and under diagnosis) to impact on health economics from inappropriate over usage of diagnostic tests. For these reasons, there have been a number of papers recommending the use of simulation for cardiovascular examinations (Perlini et al., 2014; Simon et al., 2012). Previous studies comparing the use of simulation to real patients for training of cardiovascular examination skills have predominantly focused on auscultatory training using multimedia (such as standalone digital audio and CD-ROM) and whole system examination using human manikins, though these do not simulate a key aspect of the examination - patient interaction (McKinney et al., 2013). By offering a means for more integrated training and assessment of technical and non-technical skills, PFHS may have offer potential benefits in this area. Mangione et al (1997) demonstrated in a multi-centre multinational study that internal medicine residents' ability to correctly diagnose on cardiac auscultation was low across three different countries and accuracy ranged from 20-26%. Similarly, another study found that diagnostic accuracy of cardiovascular examination amongst paediatric residents was only 30% (Dhuper et al 2007).

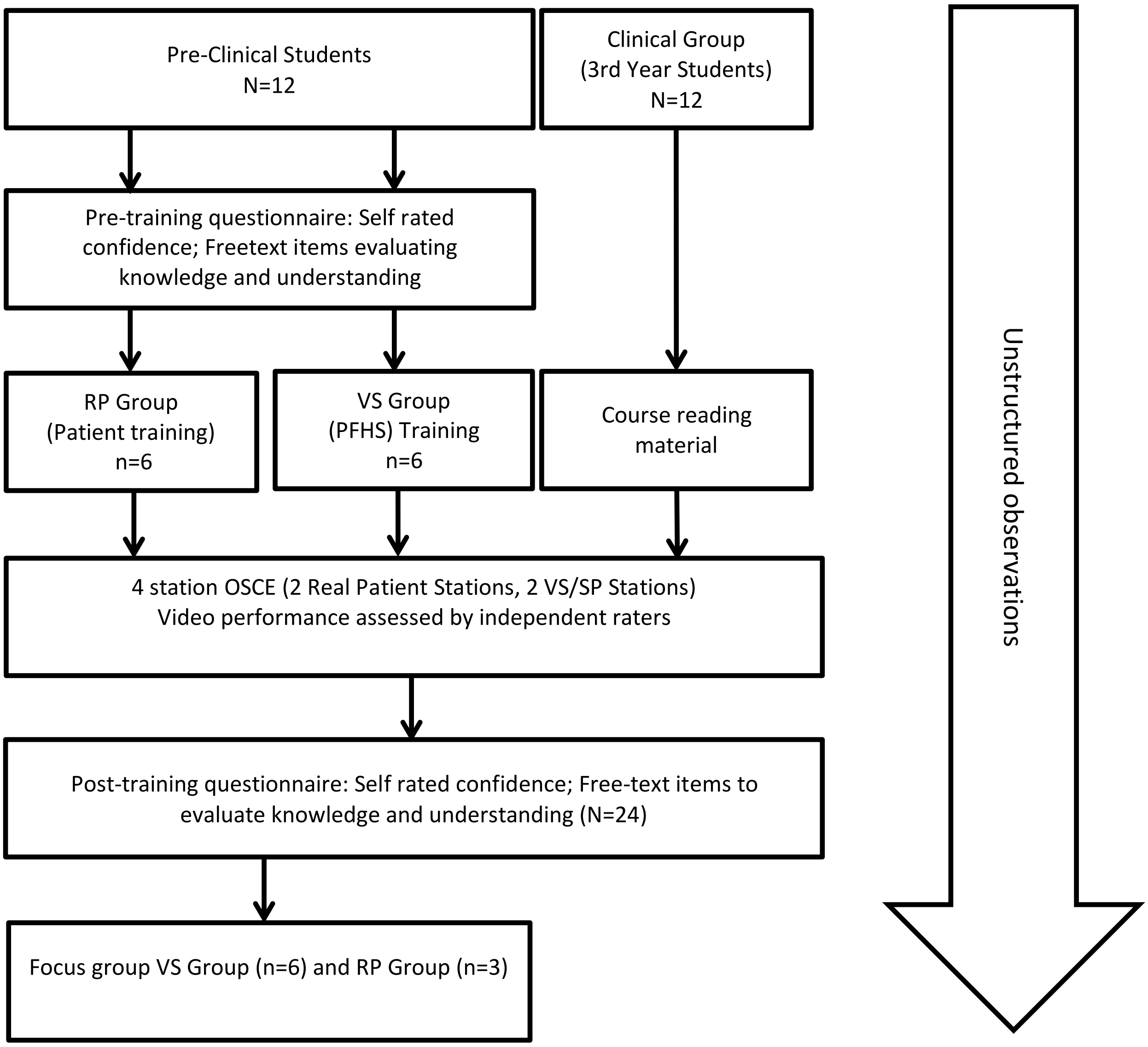
There are also several practical reasons for why I chose to study learning and assessment of cardiovascular examination skills to investigate PFHS. In this study, I am specifically interested in comparing clinical skills training with PFHS to training with real patients. From a practical and ethical perspective, it is difficult to justify a study that compares the training of invasive procedures on real patients to simulation, particularly as there is the potential to cause harm to patients. At our institution like with many other medical schools and teaching hospitals, we also have ready access to regular patient volunteers with cardiac signs that are invited to participate in training and assessment activities. With respect to creating a PFHS for cardiovascular examination skills training and assessment, we also had ready access to equipment and personnel required to create the simulations, which are already used at our institute. The specific PFHS used in this study is a hybrid of Ventriloscope® and Simulated Patient (VS/SP) which is described in more detail below.

6.2 Methods and Materials

In order to answer the key questions I have employed a mix-methods approach to the inquiry for the reasons stated in Chapter 2. Specifically, for this empirical study, the aims of my inquiry are to not only look at outcomes of learning and assessment of cardiovascular examination skills with PFHS and real patients, but also explore the learning process in order to understand mechanisms, causality and effect.

An overall map of the study protocol is shown in Figure 6.1. The individual sections of the study components will be discussed in more detail within the respective sections.

Figure 6.1 Overall Map of Study. VS = Ventriloscope®; RP = Real Patient; VS/SP=Ventriloscope®/Simulated Patient Hybrid Simulations.



6.2.1 Study setting

The study was conducted in a clinical skills laboratory of a central London medical school Ethical approval for the study was granted by Imperial College Research Ethics Committee (ICREC_9_4_1).

6.2.2 Participants:

For this study, three training interventions were compared: focused short training course using PFHS; a parallel focused short training course using real patients; and naturalistic learning in the workplace. For these reasons, students invited into the study were as follows:

- 1st and 2nd year (pre-clinical) medical students from a single medical school with no prior experience of clinical examination to control for possible effect of prior training.
- 3rd year medical students (Clinical Group) who have almost completed 1 academic year of clinical training.

Inclusion Criteria for Pre-Clinical:

- 1st or 2nd year medical students

Exclusion Criteria for Pre-Clinical:

- Prior experience of performing cardiovascular examination

Inclusion Criteria for Clinical Students

- Students who have had ward-based training in cardiovascular examination skills

Exclusion Criteria for Clinical Students:

- No prior clinical experience of performing cardiovascular examination on real patients.
- Students who have completed less than 6 months of their 3rd year clinical clerkship.
- Students with more than 1 academic year's experience of ward-based learning of clinical examination skills.

Pre-clinical students were randomised using random number allocation into two training intervention groups:

- VS Group - trained with VS/SP hybrid simulations
- RP Group - trained with patient volunteers.

Students were invited on a voluntary basis via electronic mail and telephone and provided full written consent for study participation (Appendix 7).

6.2.3 Training Interventions

All pre-clinical students (VS and RP groups) underwent a focused training course for the purpose of the study. Although recent research have highlighted the relative benefits of distributed training (over a long period of time) over focused training (short courses), for research and practical reasons, a two day course was designed to train the pre-clinical students in cardiovascular examination. A training course distributed throughout a longer period may be more susceptible to external confounding factors and educational influences. The course was designed such that the training undertaken by VS and RP group students was identical except for the practical component whereby students received PFHS or real patient training respectively. The days of the courses for each group were run in a staggered manner over 4 days to minimise difference in training to assessment time between the two groups.

Table 6.1 Cardiovascular Examinations Skills Course Outline and Schedule for RP and VS Group Students

VS Group	RP Group
Day 1 (Instruction lead by JZ)	***Day 1 (Instruction lead by JZ)***
Series of Small group tutorial	Series of Small group tutorial
Practice healthy volunteer subjects	Practice healthy volunteer subjects
Day 2 (Instruction lead by MW)	***Day 2 (Instruction lead by MW)***
Small group tutorial and review of course content	Small group tutorial and review of course content
Practice on SP with Ventriloscope® simulating aortic stenosis	Practice on RP with aortic stenosis
Practice on SP with Ventriloscope® simulating mitral regurgitation	Practice on RP with mitral regurgitation
Day 3	***Day 3***
OSCE	OSCE

The course consisted of a series of interactive small group tutorials lead by the respective tutors. A freely accessible website containing an extensive library of visual and audio multimedia instructional material was also used (www.blaufuss.org). Course material is available for review in Appendix 8.

The practical sessions involved training of cardiovascular examination skills first on volunteer medical students, then on VS/SP hybrid simulations and real patients with known underlying cardiac pathology for the respective groups under the guidance of the tutor.

Two external tutors were recruited to teach on the training course:

- MW - a fully qualified general practitioner (GP) and GP trainer.
- JZ - a senior cardiology registrar who holds the Member of the Royal College of Physicians (MRCP), UK diploma.

Both tutors had proven competence in conducting cardiovascular examinations and have extensive teaching experience at both undergraduate and postgraduate level.

The Clinical Group (3rd year students) received the course written material including the cardiovascular examination protocol (Appendix 8), but no additional training.

6.2.4 Post-Course Performance Assessment

All students were assessed for competence in a 4 station OSCE consisting of two VS and two RP Stations to increase sampling of performance. The OSCE was conducted in the typical manner in the form of a circuit (Boursicot and Roberts, 2005). Students were required to complete all 4 cardiovascular examination stations in sequence, alternating between PFHS and RP stations. Each station had a time limit of 8 minutes. Written briefings for each examination station were provided.

It is well documented that performance in OSCEs may be influenced by "exam technique"(Cooper et al., 2012; Rashid et al., 2011). Students were therefore randomised to start on either VP/SP stations to compensate for potential learning effect between stations, where their performance may improve or even worsen in subsequent stations. One member of faculty was present in each OSCE station to act as invigilator. For VS/SP stations, the trained faculty member in situ also controlled the Ventriloscope® device. Prior to the OSCE, all students in the RP and Clinical Group (3rd year students) participated in a brief familiarisation session with the Ventriloscope® device. Students in the VS group already had experience with a real stethoscope during the training course.

6.2.5 Volunteer Patient Participants

Four patients were invited via contacts from the medical school to assist with the study. The patients had a definitive diagnosis of valvular heart disease of either aortic stenosis or mitral regurgitation with corresponding audible systolic murmurs, though they were not undergoing any active medical treatment relating to these. Murmurs had to be at least Grade 3 or above in terms of audibility, i.e. moderately loud (Gaskin et al., 2000; Thompson et al., 2001). The reason for using patients with these pathologies and systolic murmurs is that it is what a medical student at our institute would be expected to detect and describe after one academic year of clinical training (end of 3rd year). The volunteers recruited did not have other significant cardiovascular clinical signs in order to match the VS/SP simulations which themselves only replicated auscultatory signs and no other clinical signs. Two volunteers, one with mitral regurgitation and the other with aortic stenosis participated in the practical sessions during the training course. A further two new volunteers were invited for the OSCE assessment.

6.2.6 VS/SP Simulation Design

The PFHS simulations for cardiovascular examination consisted of an SP and Ventriloscope®. The Ventriloscope ® is a part-task trainer designed for simulation-based training and assessment of auscultatory skills. Physically, it resembles a stethoscope and consists of 3 main components (Figure 6.2).

1) Modified Stethoscope - this is a standard stethoscope that has been modified to incorporate an audio receiver and player that can playback and transmit pre-recorded auscultatory sounds into the stethoscope earpiece.
2) Sound Transmitter - this component transmits pre-recorded audio sounds, stored in digital format on Secured Digital (SD) memory cards to the audio receiver on the modified stethoscope - There are a series of buttons which are programmed to playback the different pre-recorded auscultatory sounds.
3) Heart rate monitor - this is worn by the SP around their chest and detects their heart rate and transmits this to the sound transmitter.

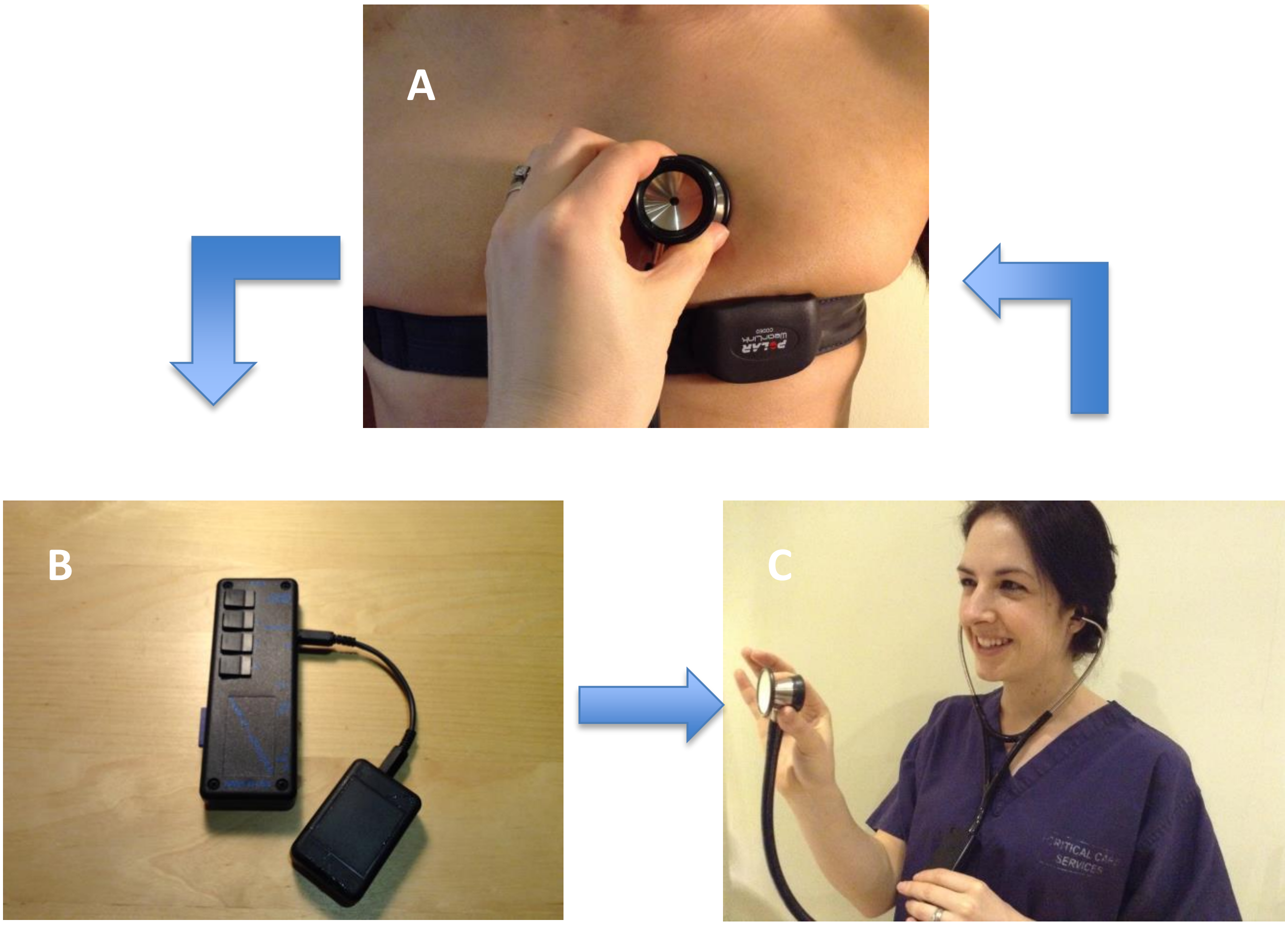


Figure 6.2 The Ventriloscope® and Main Components

The heart-rate monitor (A) sends a signal to the receiver (B,) which dictates synchronises sound playback to pulse rate. When the stethoscope is applied to the chest (A), the instructor activates the sound transmitter (B) which sends the auscultatory sound to the Ventriloscope® (C).

The Ventriloscope® can be used to create PFHS when used in conjunction with an SP. The SP provides the medium to simulate the patient-clinician communication, inspection and palpation, whilst the Ventriloscope® provides the auscultatory cues. To accurately simulate the auscultation aspect of the cardiovascular examination, an operator (usually faculty) activates the appropriate sounds when the Ventriloscope® is correctly positioned on the patient. The rate of playback of the auscultatory sounds is synchronised with the SP's heart rate, which is measured by the heart rate monitor.

The SPs used in this study are professionally trained actors who are employed on an ad-hoc basis by the medical school. The VS/SP simulations were designed to match the real patients. The SPs were instructed to behave as a clinically stable patient in an outpatient setting who is fluent in the English language and complies with instructions for clinical manoeuvres. The Ventriloscope® was set to playback either normal heart sounds, aortic stenosis or mitral regurgitation to match the cardiac pathologies of the patient volunteers. Different sound libraries for the auscultatory sounds were used in training and in the OSCEs, i.e. the aortic stenosis sounds used in the training course were different from the ones used in the assessment.

6.2.7 Learning outcomes of VS/SP trained and Real Patient Trained Students

This component of the study was to investigate key questions 2A and 2B. For the reasons discussed in Chapter 3 and Chapter 5, a mixed methods approach combining quantitative and qualitative elements was used in this study. Three key aspects were studied with respect to learning outcomes as follows:

6.2.7.1 Post-Training Cardiovascular Examination OSCE Performance

The aim of this component of the study was to:

- Compare performance of cardiovascular examination between VS, RP and Clinical group students.
- Compare students' cardiovascular examination performance in VP/SP OSCE stations to performance in real patients OSCE stations.

A video of the students' performance in each OSCE station was recorded using the Scotia Medical Observation and Training System (TM) for assessment. The videos were reviewed remotely and rated by two blinded assessors (GC a senior cardiology registrar and LSPM a senior medical registrar) independently. Performance was rated using an OSCE marking tool containing a series of checklist items (maximum score = 44) and single global rating scale for overall performance (maximum score = 6) developed by MG and JKT. The assessment tool was checked for content validity by the course tutors (JZ and MW). OSCE mark sheets are available in Appendix 9.

Kruskall-Wallis test was used to determine interaction between OSCE performance scores of the different groups of students. Spearman's rank correlation was used to determine correlation between students' OSCE scores of VS/SP and RP stations.

6.2.7.2 Self-rated confidence and competence in performing cardiovascular examination

The aim of this component of the study was to:

- Evaluate and compare students' self-rated confidence and competence in performing cardiovascular examination in RP and VS groups before and after post-training and
- Compare students' self-rated competence and confidence in performing cardiovascular examination between RP and VS groups and Clinical group post-training.

Self-rated confidence and competence was measured using a pre and post-training questionnaire which consisted of a series of 5-point Likert-type items (Appendix 9)

6.2.8 Post-Course Focus Groups

As with the earlier study conducted in Chapter 5, the underpinning theoretical foundations to this qualitative component of the study is Interpretative Phenomenological Analysis to explore the students' experience of training with VS/SP simulation. The method of data collection was through focussed groups. Focus groups are not merely group interviews, where the researcher asks the group members questions, but its power lies in the shared characteristics of its members and the dialogue between them (Kitzinger, 1995). The main reason for using focus groups was to encourage discussion amongst students, which is particularly important in this study due to the possible hierarchical tensions between the researchers, i.e. myself (as a qualified doctor and course organiser), and the participants (medical students who are early in their training). Given its nature i.e. a "group", there are a number of authors that argue that focus groups are incompatible with the Interpretative Phenomenological Analysis approach as the principle of idiography cannot be applied. Others including myself however argue that individual voices can still be "heard" so long as the moderators are mindful of this potential issue (Smith et al., 2009, chap. 4).

The focus groups were facilitated by a lead moderator (JKT) and assisted by a member of the research group (MW) and conducted in a meeting room, immediately after the OSCE. The role of the moderators was to ensure not only to present the focus group with questions in line with the research aims, but to encourage a balanced discussion and prevent certain members from being overly dominating discussions {ref}.

The focus groups lasted for 45min to 1 hour and were recorded on a digital audio recorder. The focus group topic schedule was as follows:

- General perception of the VS/SP simulation
- Learning with the VS/SP simulation
- Performing on real patients.
- Training issues
- Technical issues

6.2.9 Data Analyses

Quantitative data analysis was conducted using SPSS 20. Where statistical significance needed to be determined, an alpha level of 0.05 was used. The specific analyses for each component of the study are described within the respective sections.

Coding of focus groups was conducted by two researchers (JKT and MG) collaboratively to bring insight from two potentially different perspectives into the data set, i.e. that of a qualified doctor (JKT) and a medical student (MG). Initial open coding resulted in emergent themes that were subsequently distilled into the final superordinate and subthemes.

6.3 Results

A total of 24 students participated in the study (after 4 dropouts from the RP Group and 4 dropouts in the VS Group prior to start of course). Student background demographics are presented in Table 6.2.

Table 6.2 Demographics of Study Participants

Demographics	RP Group n=6	VS Group n=6	Clinical Group n=12
Year of study	Year 1 = 4, Year 2 = 2	Year 1 = 5, Year 2 = 1	Year 3 = 12
Mean age (SD)	19 (1.10)	20 (2.53)	21 (0.60)
Male (%)	2 (33.3%)	4 (66.7%)	4 (33.3%)
Previous Experience of:			
Learning with simulated patients	6 (100.0%)	6 (100.0%)	12 (100.0%)
Interacting with patients	4 (66.7%)	6 (100.0%)	12 (100.0%)
Examining patients	1 (16.7%)	33.3	12 (100.0%)
Conducting a Cardiovascular Examination on Patient	0 (0.0%)	0 (0.0%)	12 (100.0%)

6.3.1 OSCE Performance

Results of the students' overall OSCE performance are summarised in Table 6.3.

Boxplots of students scores for the different stations are shown in Figure 6.3 and 6.4.

Table 6.3 Students' OSCE Checklist and GRS scores for the VS/SP and Real Patient Assessment Stations

	Checklist scores			GRS Scores		
	25th Percentile	Median	75th Percentile	25th Percentile	Median	75th Percentile
VS/SP Station						
VS Group	33.9375	35.6250	37.25	3.125	3.75	4.0625
RP Group	33.25	33.625	36.75	3	4.125	4.5
Clinical Group	30.5	34.375	35.125	3.0625	3.625	4
Real Patient Station						
VS Group	32.1875	35.125	38.1875	2.6875	3.6250	4.75
RP Group	31.43	34.875	36.375	3.3125	3.8750	4.8125
Clinical Group	30.50	32.8750	33.8750	3.25	3.5	4.1875

Table 6.4 Sub-analysis of Diagnostic Performance

	Diagnostic Score		
	Diagnostic Score		
	25th Percentile	Median	75th Percentile
VS/SP Station			
VS Group	0.75	2.5	4
RP Group	1	3.5	4
Clinical Group	1	2.5	4
Real Patient Station			
VS Group	2	2	4
RP Group	1.25	2	4
Clinical Group	1	2	3

Figure 6.3 Box Plots of OSCE Checklist Performance Ratings for RP, VS and Clinical Students in VS/SP and Real Patient Stations

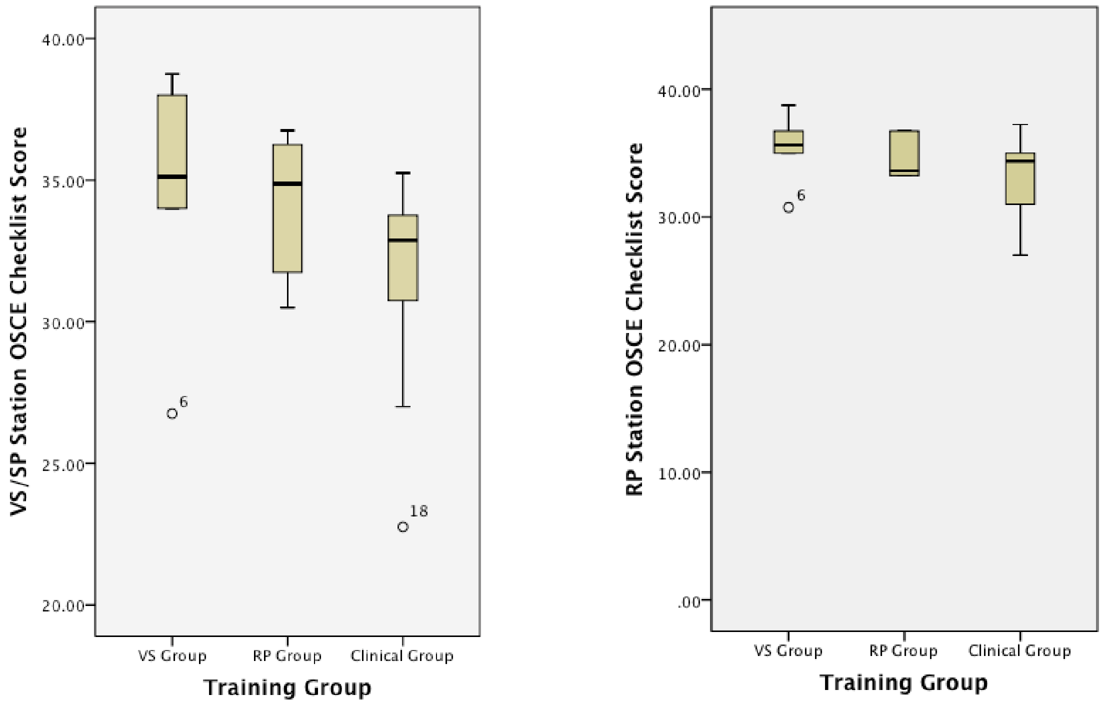
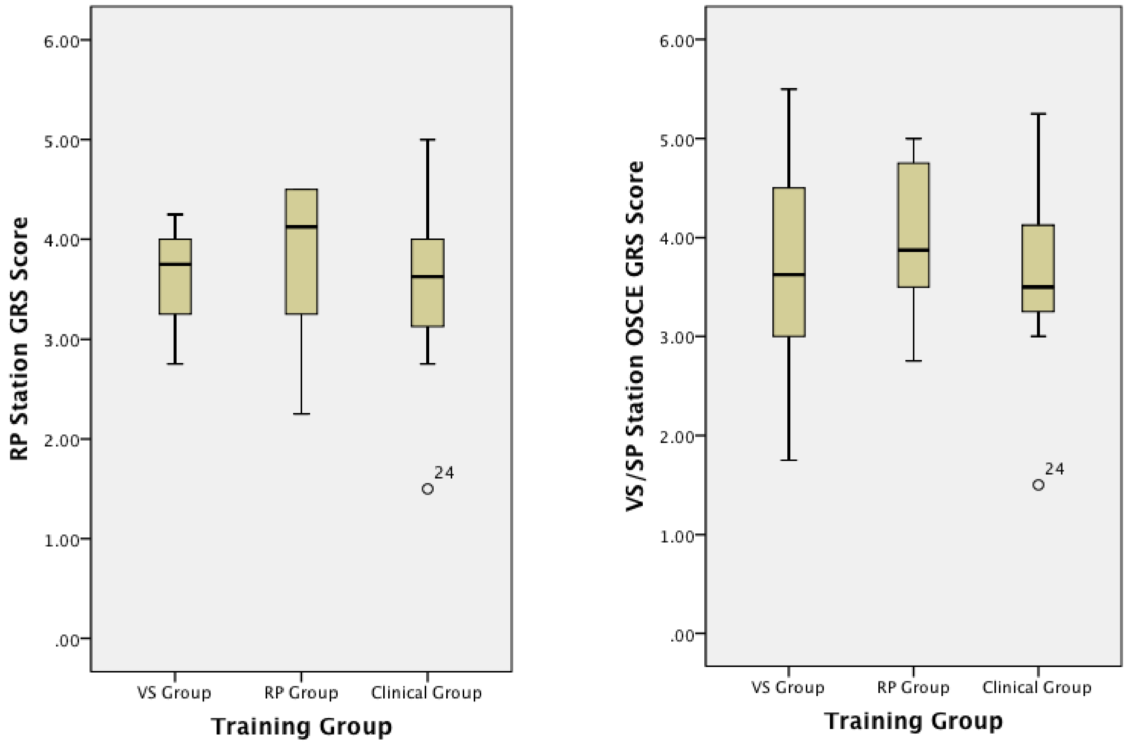


Figure 6.4 Box Plots of OSCE GRS Performance Ratings for RP, VS and Clinical Students in VS/SP and Real Patient Stations



Comparison of OSCE performance

Kruskal-Wallis analysis of variance was conducted to determine effect of type of training (VS, RP or CT) on OSCE performance. No significant effect on students' OSCE performance was detected in the real patient stations when measured by both the checklist (p=0.207) and GRS (p= 0.721). No significant effect on students' OSCE performance was detected in the VS/SP stations when measured by both the checklist (p=0.113) and GRS (p= 0.695).

Kruskal-Wallis analysis of variance was conducted to determine effect of training group (VS, RP or CT) on diagnostic performance (maximum score = 4). No significant effect of training group on students' diagnostic performance was detected in the real patient stations (p=0.636) or VS/SP stations (p= 0.353).

6.3.1.1 Correlation between OSCE Scores in PFHS and Real Patient stations

Scatterplots comparing students' scores in VS/SP and RP stations are shown in Figures 6.5 and 6.6.

Figure 6.5 Scatterplots of OSCE Checklist Scores in RP and VS/SP Stations

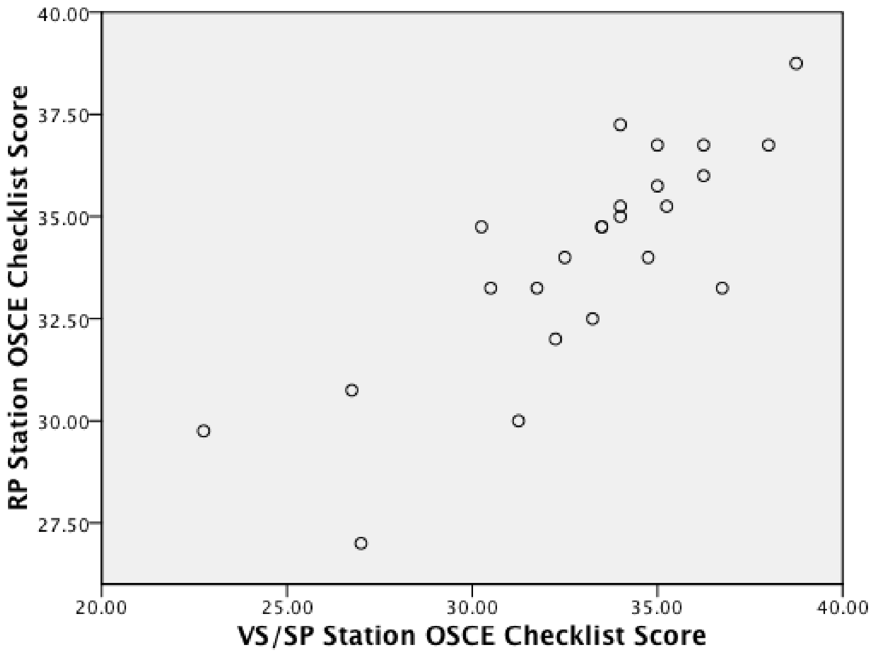
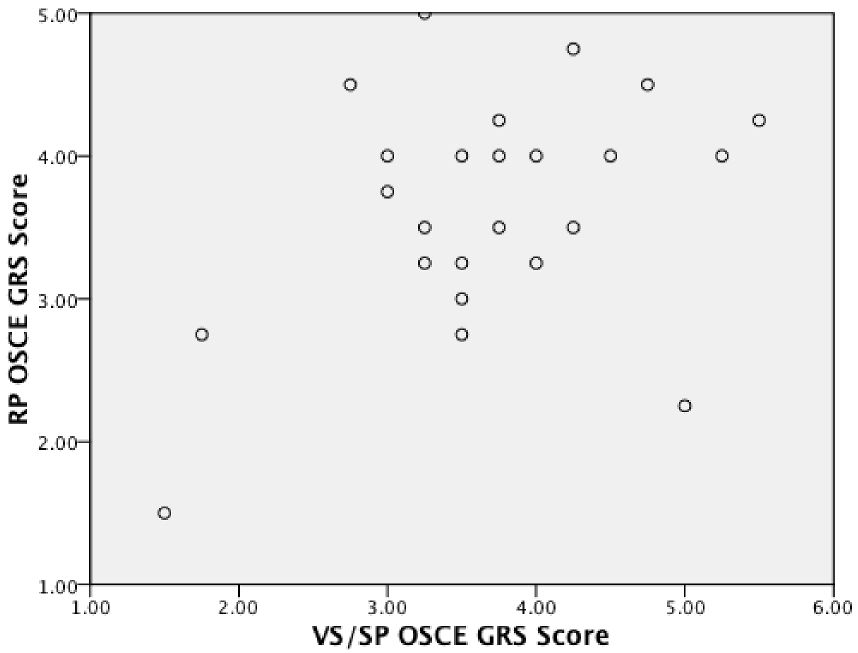


Figure 6.6 Scatterplots of OSCE GRS Scores in RP and VS/SP Stations



There is statistically significant correlation between student's OSCE checklist scores in the VS/SP station and RP stations (Spearman's ρ = 0.790, p=0.00). However, there is positive but non-statistically significant correlation between student's OSCE checklist scores in the VS/SP station and RP stations (Spearman's rho = 0.265, p=0.210).

6.3.2 Self-rated confidence and competence

Results for self-ratings in competence and confidence in performing cardiovascular examination for RP and VS Group students, with comparisons of ratings pre and post-course is presented in Table 6.5. Pre and post-course ratings were compared using Wilcoxon Exact 2-Sample Test. For VS Group Students, there is a statistically significant increase in ratings for Question Items 4, 5 and 6 and non-statistically significant increase in Question items 1, 2 and 3. For RP Group Students, there is a statistically significant increase in ratings for Question Items 4, 5 and 6 and non-statistically significant increase in Question items 1, 2 and 3.

Results for self-ratings in competence and confidence for Clinical Group is presented in Table 6.6. A list of the questionnaire items is presented in Box 6.1. Comparison of self-ratings between VS Group (both pre and post-course), RP Group (both pre and post-course) and Clinical Group was performed using Mann-Whitney U Test and results are presented in Table 6.7. There was no statistical difference in pre-course self-ratings for all question items between VS and RP Groups. There was also no statistical difference in post-course self-ratings for all question items between VS and RP Groups. There was no statistical difference in post-course self-ratings for all question items between Clinical and RP Groups, and between Clinical and VS Groups.

BOX 6.1 Self-rated Confidence and Competence Questionnaire Items

Q1) I feel confidence in having direct physical contact with patients for the purposes of clinical examination

Q2) I feel confidence in communicating with patients

Q3) I feel confident in using a stethoscope

Q4) I feel confident in conducting a cardiovascular examination on a real patient

Q5) I feel competent in my knowledge of cardiovascular examination skills

Q6) I feel competent in conducting a basic cardiovascular examination.

Table 6.5 Self-ratings for Confidence and Competence of VS and RP Group Students in Performing Cardiovascular Examination Skills Pre and Post Course

Self-Rated Confidence and Competence in CVS Examination							
	Pre-course			Post-course			Pre vs Post-course
Questionnaire Item	25th Percentile	Median	75th Percentile	25th Percentile	Median	75th Percentile	
RP Group							
Q1	1.75	3.5	4.25	3.75	4	4.24	0.250
Q2	2.75	3.5	4	4	4.5	5	0.063
Q3	1	2	3.25	3	3	5.25	0.125
Q4	1	1	2	4	4	4.25	0.031*
Q5	1	1	1.25	2.75	3.5	5	0.031*
Q6	1	1	1	2.75	3.5	4	0.031*
VS Group							
Q1	2.75	3.5	4	4	4	5	0.125
Q2	3	4	4	4	4	4.25	0.250
Q3	1.75	2	3	3.75	4	4.25	0.063
Q4	1	2	2	3.75	4	4	0.031*
Q5	1	2	2	3	4	4	0.031*
Q6	1	2	2	3.75	4	4.25	0.031*

Table 6.6 Self-ratings for Confidence and Competence of Clinical Group Students in Performing Cardiovascular Examination Skills

	Self-Rated Confidence and Competence in CVS Examination		
Question Item	25th Percentile	Median	75th Percentile
1	4	5	5
2	4	5	5
3	4	5	5
4	4	4	4.75
5	4	4	4
6	4	4	4

Table 6.7 Comparison of self-ratings between RP and VS Groups (Pre and Post-course), RP and Clinical Groups (Post-course) and VS and Clinical Groups (Post-Course)

Item	P-values for Comparison of Self-Ratings in Confidence and Competence			
	Pre-course RP vs VS	Post-course RP vs VS	Post-Course RP vs CT	Post-Course VS vs CT
Q1	0.937	0.485	0.102	0.437
Q2	0.589	0.394	0.820	0.18
Q3	0.937	0.240	0.18	0.102
Q4	0.394	0.394	0.820	0.213
Q5	0.180	1	0.437	0.151
Q6	0.065	0.240	0.53	0.682

6.3.3 Knowledge and Understanding of Cardiovascular Examination Skills Pre and Post Course

All pre-clinical students from both the VS Group and RP Group completed the pre and post-course questionnaire (after OSCE). All students in the Clinical Group (3rd years) also completed the questionnaire after completing the OSCE. Content analysis of the free-text questionnaire responses are presented in tables 6.9, 6.10 and 6.10 for question items 1 to 3 respectively. An anonymised sample of responses is provided in Appendix 11.

Table 6.8 Content Analysis of Free-Text Responses to Knowledge and Understanding of Cardiovascular Examination Questionnaire (Q1 What is the purpose of the cardiovascular examination?)

	Incidence of Response to Question (%)				
	VS (n=6)		RP (n=6)		3rd Year (n=12)
Response Content	Pre-Course	Post-Course	Pre-Course	Post-Course	Post-Course
General cardiac function	3 (50%)	6 (100%)	4 (66.7%)	2 (33.3%)	3 (25%)
Identifying pathology	3 (50%)	5 (83.3%)	3 (50%)	4 (66.7%)	11 (91.7%)
Elicit clinical signs	0 (0%)	4 (66.7%)	2 (33.3%)	6 (100%)	6 (50%)
Examining whole CVS system	0 (0%)	3 (50%)	0 (0%)	3 (50%)	0 (0%)

Types of responses and changes in responses pre and post course were similar between VS and RP training groups. Pre-course, students' perception of the purpose of the cardiovascular examination was to assess the general function of the heart. Post course, student responses were more weighted towards identifying pathology and eliciting clinical signs. In addition, there was an increased emphasis of examination of the whole "system" as opposed to just examining the heart. Sample of anonymised responses are presented in Appendix 11

Table 6.9 Content Analysis of Free-Text Responses to Knowledge and Understanding of Cardiovascular Examination Questionnaire (Q2: What are the key components of the cardiovascular system?)

	Incidence of Response to Question (%)				
	VS (n=6)		**RP (n=6)**		**3rd Year (n=12)**
Response Content	**Pre-Course**	**Post-Course**	**Pre-Course**	**Post-Course**	**Post-Course**
Inspection	2 (33.3.%)	5 (83.3%)	0 (0%)	6 (100%)	10 (83.3%)
Palpation	3 (50%)	6 (100%)	0 (0%)	6 (100%)	11 (91.7%)
Auscultation	3 (50%)	5 (83.3%)	5 (83.3%)	6 (100%)	12 (100%)
Examining whole CVS system	0 (0%)	2 (33.3%)	0 (0%)	2 (33.3%)	5 (41.7%)
Communication	0 (0%)	1 (16.7%)	0 (0%)	1 (16.7%)	0 (0%)

Types of responses and changes in responses pre and post course were similar between VS and RP training groups. Pre-course most students stated auscultation as a key skill required in cardiovascular examination, though few stated other aspects of the cardiovascular examination, i.e. inspection and palpation. Post-course, almost all students stated inspection, palpation and auscultation as key components of the cardiovascular examination.

Table 6.10 Content Analysis of Free-Text Responses to Knowledge and Understanding of Cardiovascular Examination Questionnaire (Q3: What are the key skills required to perform a cardiovascular examination competently?)

	Incidence of Response to Question (%)				
	VS (n=6)		RP (n=6)		3rd Year (n=12)
Response Content	**Pre-Course**	**Post-Course**	**Pre-Course**	**Post-Course**	**Post-Course**
Core Knowledge	2 (33.3%)	4 (66.7%)	3 (50%)	3 (50%)	7 (58.3%)
Technical	4 (66.7%)	4 (66.7%)	3 (50%)	5 (83.3%)	8 (66.7%)
Communication and Professionalism	4 (66.7%)	4 (66.7%)	5 (83.3%)	5 (83.3%)	8 (66.7%)
Fluency	0 (0%)	2 (33.3%)	0 (0%)	0 (0%)	2 (16.7 %)

Types of responses were similar between VS and RP training groups. Students' pre-course responses were similar to post-course responses. Students stated the need for knowledge and technical skills. In addition, students stated the need for communication skills and professionalism. Throughout the responses of the three questions there was a distinct change in language from more layman-like to "medical". For example, students tended to used terms such as "look", "feel" and "listen" pre-course, and "inspection" palpation" and "auscultation" post course. Following training, some students stated the need to perform the examination with fluency. One student specifically commented on the need to be able to perform under time pressure.

6.3.4 Results of Focus Groups

Two focus groups were conducted in series. Nine students participated in the focus groups in total, all 6 students from the VS group and 3 students from the RP. Thematic analysis of the focus group dialogues gave rise to three major superordinate themes - *Learning with VS/SP*, *Transfer to the real patient*, and *Course Issues* each with some several associated subthemes. An overview of the themes is presented in Table 6.11

Table 6.11: Overview of Themes from Analysis of Focus Groups

Superordinate theme	Subtheme
Learning with the VS/SP Simulation	Developing a foundation
	Integration of knowledge and skills
	Importance of focused learning of component skills
	Limitations of using real patients for training
	Technological issues of VS
Performing on real patients	Level of confidence
	More difficult on performing on real patient
	Margin of error in simulation
	Importance of practicing on real patients
Course Issues	Lack of non-auscultatory clinical signs
	Information Overload

6.3.4.1 Learning with the VS/SP Simulation

Several sub-themes were identified which provided insight into the process of learning with the VS/SP hybrid in the context of this short course.

Developing a Foundation

Students commented on the usefulness of VS/SP as a means of developing a foundation for real clinical practice. Key to this is the scope for developing a routine, particularly through repetition and automation of the examination process. The process of developing a routine and fluency of performing the cardiovascular examination allowed students to concentrate on interpreting the signs and identifying pathology.

> *"… you've got to have some sort of a foundation to build upon and the Ventriloscope gives you that foundation very clearly and then you obviously move on to the next stage which is building up the experience and that is when the real patient comes in" (VS4)*

"By the end of the course I am thinking much much less about structure, but looking back at the very first day (I was wondering) how am I going to remember all the components and already … by today I was thinking more of what is the diagnosis rather than what is the next step… getting used to the routines and getting your skills fluent… (VS1)

Integration of knowledge and skills

Another key process in the acquisition of the cardiovascular examination skills with the VS/SP was the ability to allow students to apply and integrate their knowledge and skills in a practical experience. Students specifically commented on the value of learning various components of the cardiovascular examination in context during the simulation, which allowed them to make sense of the theory.

"…the multimedia (digital media of auscultatory sounds in isolation) *was very clear and good, but it was completely out of context listening to it on over speakers and when you do it with the Ventriloscope you have the stethoscope in different positions, used the bell on the diaphragm and you hear the sounds more in context to where they would be." (VS5)*

"…because we had the practical (VS/SP simulation) *soon afterwards* (after the tutorial) *, we could put it into context and I would be seeing and looking for things that I remember from the lecture when we doing the examinations. I think it has helped to connect everything quite well." (VS5)*

Importance of focused learning of component skills

In this training course, the VS/SP PFHS was designed to provide students a medium to practice their skills in an integrated fashion, combining both technical and non-technical aspects, however, some students stated the importance of focused learning of component skills and tasks.

"…it was nice having the Ventriloscope to have a really good focus on just listening to heart sounds…I do think you do feel a lot more confident once you start with the

Ventriloscope and essentially is more accessible … it's focusing on a single symptom as opposed to the patient" (VS5)

Limitations of using real patients for training

A number of students suggested that the VS/SP simulation could be used to overcome some limitations of using real patients for training, specifically in relation to access and the practicalities of organising such a course.

> *" Especially for this much accelerated kind course. I think it's much easier for the people organising it as opposed to getting all sorts of different patients in order to keep within schedule. You can* (also) *have more choice* (of pathology) *then." (VS3)*

Technological issues

A number of technological issues were identified by the students with respect to the use of the Ventriloscope®. These included: a slight delay between activation of the Ventriloscope® by the operator and the transmission of sound; occasional failure of the Ventriloscope® in terms of producing sounds; and the presence of the heart rate monitor worn around the SP's chest which was perceived to be obtrusive particularly when students were palpating the chest. A number of students also commented on the need to familiarise themselves with the technology.

> *"… the Ventriloscope has certain technical issues which make gave it irregular rhythms and was kind of hard to hear." (VS2)*
>
> *" One of the things that was disconcerting about the Ventriloscope was that sometimes you put it on there* (patient's chest) *for two seconds and it was three seconds before actually started to hear things." (VS4)*
>
> *"…when you are palpating for an apex … because there are physically things* (heart rate band) *in the way (VS6)*

6.3.4.2 Performing on real patients

Level of Confidence

The majority of students agreed that they would feel confident in performing a cardiovascular examination on a real patient after training with the VS/SP in the course, though they stated that this would ideally be under supervision. The level of confidence in performing the cardiovascular examination on a real patient was also variable between students, and in particular with respect to the different components of the cardiovascular examination. For instance, certain aspects of the cardiovascular examination were recognised by student to be challenging, such as detecting and interpreting the different auscultatory sounds.

> *"I feel confident to conduct it. I think I know what I am supposed to do and I feel like it is relatively fluid and moving from each step, but I don't feel confident in being able, I mean I know what I am supposed to look for, I know it radiates to the axilla, for instance, and it is supposed to be MR (mitral regurgitation), but I wouldn't feel confident in my ability to detect the sounds." (VS6)*

An area which students appeared to have confidence in following simulation training is in interacting with real patients, possibly due to the perceived resemblance of interacting with an SP to with a real patient.

> *"I definitely feel more confident with talking to patients as well; you know communication as well and knowing to deal with a patient. That* (VS/SP) *has definitely helped I think"*
> *"To be honest, we actually had an actor acting out to be a patient is only a matter of the sound using the stethoscope, so when I actually went in to see a patient it was just like that. No different at all." (VS1)*

More difficult on performing on real patient

Students in general perceived that performing cardiovascular examination on a real patient presented a greater level of difficulty in comparison to performing on a VS/SP simulation. A major factor in the difference in challenge was the relative ease of listening to and interpreting the auscultatory sounds on the Ventriloscope device, which was clearer. Auscultatory sounds also varied from real patient to another, which provided additional difficulty.

> *" I think the Ventriloscope does have some slight issues in that it is bit too easy to hear" (VS2)*

> *"Well listening to the sounds, sometimes the murmurs themselves weren't as obvious* (in real patients) *and weren't as pronounced, so it was hard to work out where they were in a heartbeat and everything. (VS3)*

> *" …(in) the real patient, well there is a lot of variation …. the Ventriloscope is always the same…" (VS1)*

Margin of error in simulation

A possible factor in the VS/SP simulation which may have contributed to its relative ease in terms of auscultation is the greater margin of error allowed in the positioning of the stethoscope. Where as on the real patient, the student must precisely position the stethoscope in order to obtain good auscultatory sounds, with the VS/SP, the sounds are played when the sound operator thinks that the student has positioned the stethoscope in the appropriate area.

> *"A little bit artificial because obviously you are placing the stethoscope in the position. I mean, you might find the apex, forget where the apex is and then place the bell somewhere else, but it would still deliver the sounds if the technician thinks that you are in the right kind of area … On a real patient you are going to hear nothing..." (VS6)*

The greater margin of error that was allowed with the VS/SP simulation appeared to also create the potential for negative learning. As illustrated from this student's experience during training, the sound operator occasionally activated the sounds regardless of whether the student was using the bell or diaphragm component of the stethoscope, thus the student was prone to forgetting to switch between the two (bell / diaphragm) when performing on the real patient.

> *"…when you practice a lot with the Ventriloscope and then you move on to real patients it is quite easy to forget to turn it (bell/ diaphragm), because you haven't been doing it* (during training)*." (VS1)*

Importance of practicing on real patients

Students commented on the importance of practicing on real patients to achieve competence in cardiovascular examination.

> *"I think you can use a Ventriloscope, but it is not strictly necessary for me. Eventually you have to use patients. I think it is just the experience of the patients that's important" (VS4)*

6.3.4.3 Course Issues

Students generally commented positively about the course in terms of the learning experience and value. However, they also commented on a number of issues related to the design of the training course and content, which provided insight into the learning with VS/SP in this context.

Lack of non-auscultatory clinical signs

With respect to the simulations, students commented on the lack of clinical signs other than the auscultatory signs reproduced by the Ventriloscope® in the VS/SP simulations. In particular, they noted that this lack of clinical signs in the simulations made it difficult to identify signs when they examined the real patients for the first time. Students in the RP training group also noted the lack of non-auscultatory clinical signs in the volunteer patients in the course. A potential effect of this lack of certain clinical signs during the training is that students appeared to skim pass part of the cardiovascular examination (inspection, palpation etc.) with no positive clinical signs in order to progress the auscultation component, where they were certain pathological signs may be detected.

> *"Because we had done it so much with people that looked perfectly normal and it is only that when you get to the listening, that we start really tuning and thinking there is going to be something. So when it came to the examination* (OSCE assessment) *it was very much* (a matter of) *do it quickly and get listening." (VS1)*

> *"So probably in terms of peripheral signs, checking to the things you see in the hands, it probably does not matter" (VS3)*

Interestingly, the students were not always aware of when they missed a sign as highlighted by this dialogue. This may be been a result of the lack of non-auscultatory clinical signs during training.

> *"Today when I saw clubbing of the fingers I wasn't quite sure if it was clubbing of the fingers or something else. So it is quite a separate skill that you need to acquire throughout the training then the skills you need when you are listening to the heart. (VS2) I actually missed the clubbing of the fingers. I don't know why. (VS4) I didn't know if I actually got it (VS3)*

> *"…because I had gotten into such a routine it was very much (take) a look at the chest but appears to be nothing obvious and didn't know how to look. So I didn't notice." (VS4)*

Information Overload

An issue raised by students was the course intensity resulting in a degree of information overload. Also, much of the course was designed to prepare students in the practical skills of performing a cardiovascular examination. However, students commented on the need to have also have adequate and integrated training in the pathological and theoretical underpinnings of the cardiovascular examination.

> *"I think going back to my point about the fact that the course /tutorials were intense) I think what might be better was to have put the pathology bit of the various peripheral signs at the end maybe? At least have a recap at the end." (VS4)*

> *"I think you did a lot of sign work well, obviously it is clinical skills but within a lot of the local signs…I think it would be nice to just to have a recap of actual physiological reason for the symptom." (VS2)*

6.3.5 Summary of Unstructured Observations

In this section I will present a series of findings from my unstructured observations and, where appropriate, excerpts from my field notes to illustrate themes. (Detailed account of observations available in Appendix 12).

Learning with the Ventriloscope® and RP

The process of learning with the VS/SP simulations and RP in the context of this course was similar in many ways. In essence, both the RP and VS/SP simulation allowed the students (who had no prior experience in cardiovascular examination) to apply their knowledge in the practical setting, under the guidance and instruction of an expert. Learning was reinforced by repetition and throughout the course, the level of tutor input required diminished.

The scope to repeatedly practice the entire examination was important in skill acquisition given the large number of steps and manoeuvres required. In addition to practicing the whole examination in full, students also selected key areas for more focused repetitive practice during the sessions. Towards the end of the course, some students from the VS group chose to use the Ventriloscope® in isolation, miming the positioning of the stethoscope rather than practicing it on a SP, even when the SP was "free" to practice on.

Role of the Tutor and Role Modelling

The tutors played a vital role within both the simulation not only in terms of direct instruction and providing technical medical knowledge, but also through role modelling. During the practical sessions, the tutors stressed the importance of good communication and professionalism from the outset, and continued to demonstrate professional behaviour to both the RP and SP.

Educational Agendas and Objectives

An observation from both the VS and RP courses was that there were two distinct educational agendas and objectives. On the one hand, tutors concentrated on aspects of good clinical technique:

> *"…(Remember) to turn the head only slightly, otherwise it will be difficult to see the JVP"*

On the other hand, tutors stressed the importance of good OSCE "examination technique":

> *"Introduction and washing your hands are easy points…. And you can fail the exam if you do not do these"*

> *"Always remember to thank the patient"*

Some students were concerned with being able to perform the different components of the cardiovascular examination with "perfection" with the aim of achieving high scores.

> *"Do you loose marks for not commenting on palmar erythema?"*

Surface Learning

There was some apparent surface learning observed during the training in both RP and VS groups. When practicing the routine of the cardiovascular examination, students sometimes skimmed over certain aspects particularly during the inspection component of the examination.

Limitations of using a real patient

With the RP training, as there was only one patient, the students had to queue to practice their auscultatory skills. A degree of fatigue in the patients was also observed towards the end of the practical sessions and the tutors introduced breaks for the patient.

Challenges of using the Ventriloscope®

There were some technical issues encountered when using the Ventriloscope® device – specifically, the device needed to be reset a number of times throughout the course due to technical failure, which periodically caused some interruption to the training.

6.4 Discussion

The purpose of this study was to investigate the use of PFHS for teaching and assessment of cardiovascular examination skills, comparing it to with patients. In addition to learning and assessment outcomes, a key objective was the exploring the process of learning with PFHS. The study gave rise to a number of key findings.

This study demonstrated that preclinical students with no prior experience in performing cardiovascular examinations who were trained with PFHS (VS Group) performed similarly to preclinical students who trained with real patients (RP Group). Both these groups of students also performed similarly to clinically trained students (Clinical Group) who had almost one academic year of training in the real clinical workplace. Specifically, there was no statistical difference in terms of their performance when rated using two assessment tools: OSCE checklist and GRS. Further sub analyses of the performance ratings demonstrated that there was no statistical difference between the three groups with respect to diagnostic scores. It is worth noting, however, that whilst diagnostic scores were similar between the groups, it was generally low (2-2.5 of 4). The reasons for this may be the short training time that the students (both preclinical and clinical) have undertaken. It is generally accepted that clinical examination skills take time to develop and master, particular for aspects that are more challenging such as auscultation. Studies have shown that, even for fully qualified doctors, diagnostic yield of auscultatory examination can be poor (Mangione and Nieman, 1997).

There was strong inter-rater reliability of OSCE performance ratings (Cronbach's A ranged from 0.76 to 0.85). Students' performance ratings, in terms of the OSCE checklist ratings, strongly and significantly correlated between VS and RP stations, i.e. concurrent validity was demonstrated. This is to the best of my knowledge, the first time concurrent validity of PFHS-based assessment has been demonstrated.

In terms of pre-clinical students' knowledge and understanding of cardiovascular examination, the VS and RP groups exhibited similar responses before and after the training course. Following the training course, there was a change in their knowledge and understanding. For students' understanding of the purpose of the cardiovascular examination, there was a trend for students' responses to change from evaluating "cardiac" function to assessing the "cardiovascular system" for pathology and identifying signs. In terms of students' understanding of the key components, there was a trend for students' response to change from a focus on auscultation to a more holistic process of inspection, palpation and auscultation. Students appeared to be aware of the importance of good communication and professionalism with similar responses pre and post-course. The reasons for this, maybe due to the emphasis on communication skills training, which I will discuss in more detail later. There was also a change in the language used by the students, which tended to be more technical following training. For example, pre-course students used terms such as "look" and "listen, whereas after training students tended to used the terms "inspection" and "auscultation".

A number of students commented on the need to perform under pressure and with fluency after the training. The reasons for this may be due to the nature of the course and post-course assessment, which correlates to some of the findings of the focus group discussed below. Our students were trained to not only perform cardiovascular examinations, but also perform them to the standard in order to pass an OSCE exam, which requires them to complete a thorough examination in 8 minutes.

When looking at the VS and RP students' self-rated confidence and competence in performing a cardiovascular examination, there was a lack of change in ratings for certain question items - specifically the confidence in having direct physical contact and communicating with patients. A closer look at the self-ratings demonstrates that the reasons for this may be due to a relatively high pre-course self-rating of these question items. Interestingly, as stated earlier, pre-course free-text responses of these students also demonstrated that they were aware of the need for good communication and professionalism prior to the training. In our medical school as in many other medical schools in the UK, students are exposed to real and simulated patients at an early stage of their medical training as demonstrated in the demographic data. This is largely dictated by their curriculum, which is designed in response to the recognised need to encourage development of good communications skills and professionalism early in their career. The practice of good communication skills and professionalism also appears to be reflected in the student's actual performance. Whilst I did not specifically assess the students' communication skills with a separate rating tool, the raters informally commented that many of the preclinical students demonstrated a remarkable level of communication skills considering they had never performed a cardiovascular or other physical examination on real patients.

The focus group and observational data yielded some key findings. First of all, students regarded the course as a useful learning experience. Students commented on the usefulness of VS/SP training as a means for repetitive practice, foundation building, and a platform to consolidate theoretical knowledge through application and experimenting in the simulation, which are in keeping with the learning theories of deliberate practice and experiential learning described Chapter 1. The process of learning with the VS/SP simulations and real patients appeared to be similar, which is not surprising given that the training course was designed to be almost identical. The main difference perhaps was that in the VS Group, students were able to practice and hone in on their auscultatory skills in isolation using the Ventriloscope® only without the SP. This was reflected in focus group findings, whereby students commented on the usefulness of the Ventriloscope® as a standalone tool for learning auscultatory sounds, despite being provided with a VS/SP simulation which was designed to allow integrated practice of skills. This correlated with an observation I described in the previous section, whereby towards the end of each VS/SP practical sessions, certain students voluntarily broke off to practice component parts of the cardiovascular examination such as auscultation. As described in the previous section, some students preferentially simulated positioning the Ventriloscope® in the various key cardiac auscultation areas on a table top, despite the availability of a SP. In addition to focused auscultatory practice, students also concentrated on practicing components such as palpating for the apex beat and inspecting the jugular venous pressure. The reasons for these are not certain though it may be that the students, having practiced on the full VS/SP simulation, had identified individual areas that they believed required further focused practice in the manner that Ericsson had described in his theory of deliberate practice.

These findings raise some important questions of how and when PFHS and isolated part-task trainers should be used. For instance, given that the students elected to practice their ausculatory skills with the Ventriloscope® in isolation having practiced on a VS/SP, would it be equally beneficial in terms of learning effectiveness to just expose them to real patients followed by allowing them to undertake focused practice of component parts of the examination? Of course these questions also need to be considered in the light of other aspects of educational utility such as cost, acceptability and feasibility. There are at present recognised problems with student access to suitable patients due to issues such as the current large numbers of medical students in the UK and increased acuity and severity of the current inpatient populations' illness, therefore PFHS can in theory offer some benefits in terms of access to training (Scalese et al., 2008).

In addition to the process of deliberate practice and experiential learning which rely on a student's insight and self-direction, findings from my observations also revealed the key role that the tutors played. On top of teaching, the tutors acted as important role models, demonstrating model behaviour and excellent communication skills, which may explain the general level of professionalism observed amongst students during the OSCE assessment. From the perspective of teaching and facilitation, the reliance on guidance and prompts from the tutor by the students were initially high, but this diminished towards the end of the course. Likewise, although the tutors were not briefed to do so, they appeared to progressively "let go" of their students as their skills developed throughout the course. In essence, an educational process commonly termed "scaffolding" has taken place. Scaffolding was first described by Wood et al (1976) and derived from the concept of the Zone of Proximal Development which is attributed to Vygotsky's work in the field. The Zone of Proximal development is the distance between what a learner's actual level of development is and the potential level of development in the presence of an expert. The presence of the expert facilitator can allow students to perform tasks beyond what they are capable of through guidance and supporting them in areas in which they require further support.

In the case of this study, for example, one of the tutors in recognition of the challenge that students faced when palpating the apex beat would frequently physically guide students' hands to the correct position. Scaffolding extends this concept to when tutors tailor the amount of guidance and assistance to the learner throughout their development, such that eventually the tutor's input is minimal and the learner can reach a point of independent practice (Kneebone, 2005). Based on these observations, the learning effect achieved by my student participants was dependent not only on the availability of the simulation or patient for practice, but also on a range of influencing factors including facilitation by the tutor. This raises another interesting issue in that much of the research in the current literature tend to conclude on the outcomes of simulation training in isolation. However, in reality, and in this study, there are multiple inter-related educational influences that have taken place, which may have an effect on learners' outcomes. A deeper understanding of how these various influences affect simulation training is necessary to promote best practice (Issenberg et al., 2011). Promisingly, there are at current a few studies that explore some of these other issues, such as the effect of tutor feedback, though the available evidence remains limited (Boyle et al., 2011; Issenberg et al., 2005).

There were some unexpected learning effects of the training course and assessment, which were most apparent on observation and focus group analysis. Specifically, there appeared to be evidence of some superficial learning whereby students skimmed over parts of the clinical examination such inspection. This was also evident when the students were performing in the OSCE assessment, where some students stated in the focus group interviews that they called out to the examiner what they were looking for, but did not necessarily see the signs. For instance, during the OSCE, one of the real patients had a pacemaker with a subtle scar and although many students stated they were looking for scars, many did not actually detect it.

There are several possible reasons that could explain these phenomena. First, the course was short and resulting in insufficient time for students to develop deeper learning and understanding. Secondly, as discussed in the methods section for reasons of study design, aside from the auscultatory signs, both VS/SP simulation and the real patients used in training had no other significant cardiovascular clinical signs. This may have inadvertently diverted students learning to focus on signs that are identifiable. The lack of clinical signs is a recognised problem with the use of SPs (Cleland et al., 2009). Finally, the method of assessing the students, i.e. OSCE checklists may itself be a contributing factor. Whilst there is little in the peer-reviewed literature regarding some of the side effects of the OSCE assessment format, a look through many commercially available revision guides will reveal an emphasis on tick boxing and point scoring. Many of these points are awarded simply from saying key phrases during the examination and may therefore not necessary discriminate students who have a more superficial or deeper understanding of the clinical skill being assessed. A study by Martin et al demonstrated no difference in OSCE performance between students who tended to use deep learning strategies to surface learners (Martin et al., 2000).

The focus group also revealed some interesting issues with respect to learning with Ventriloscope® itself. Whilst most students were satisfied with the technology, they generally found auscultation less challenging with Ventriloscope® when compared to on real patients. Reasons for this include comparative ease of listening to the sounds played through the Ventriloscope®. Students also commented on the greater margin or error permitted when they used the Ventriloscope®. For example the correct sounds will still be played if the person activating the Ventriloscope® sounds (faculty) perceived the students to be placing the Ventriloscope® on a reasonable anatomical location on the patient. On the other hand, poor placement of stethoscope on a real patient may result in poorer quality and non-diagnostic auscultatory sounds. Likewise students commented on how the faculty sometimes activated the sounds regardless of if the student used the bell or diaphragm end of the Ventriloscope® correctly. Consequently, there was potential risk of negative learning.

Providentially, as evidenced from the focus group findings, the students had insight into this potential learning pitfall. Despite this particular inaccuracy in the VS/SP simulation, students in the VS Group still performed equally well to the RP and Clinical Groups when performing on real patients. The reasons for this may be either that the VS/SP simulation was still sufficiently realistic to achieve a similar training effect to when training with real patients, or that the real patient OSCE assessment could not detect these errors. Regardless of the reason, there needs to be an awareness of these potential negative learning effects of PFHS and simulation in general need, which need to be acknowledged when used for training. Of note, despite the increasing peer-reviewed literature base for simulation, there are at present few available studies reporting negative aspects of simulation training, which may reflect a possible publication bias as discussed in Chapter 2 (Cook et al., 2012). Greater research effort to understand negative aspects of simulation training should be encouraged to address these issues.

6.4.1 Limitations and Considerations

There are a number of limitations and considerations that need to be taken into account when interpreting these findings. First, the number of participants was low. This was in part due to limited resources, but also unfortunately, 8 pre-clinical students (40%) dropped out immediately before the study due to time commitment issues. It was also not possible to repeat the study to increase participant numbers due to the study design - a key aspect of the study design was to compare preclinical students who had completed the short training course we provided, to students who have acquired cardiovascular examination skills in the natural setting - the study was therefore conducted at a specific time of the academic year where the students in the Clinical Group had gained almost a full academic year of workplace based experience. As a result of the low participant numbers statistical analyses were underpowered. In particular, the results that demonstrated a lack of “statistical difference” need to be interpreted with care. With respect to the qualitative aspects, the inferences that can be drawn from the results may also be less generalisable. The unstructured observational findings were derived from one researcher - myself. However, these findings were analysed concomitantly with other data from the study allowing for some degree of triangulation and verification.

Second, with respect to the performance of the students in the Clinical Group, i.e, 3rd year students, there was little data as to how much experience and exposure they had to patients during their clerkship. The nature of acquisition of clinical experience in the workplace is known to be variable and it may be that these students had not acquired sufficient experience to perform better than the pre-clinical students. Another consideration is that the students were invited to participate on a voluntary basis due to practical reasons and research ethics, which may have introduced a degree of bias in that they may exhibit different learner traits to their peers who had not volunteered.

The OSCE performance results need to be interpreted in light of the study. Although this study directly compares training and assessment of cardiovascular examination skills using PFHS with using volunteer patients, it does not compare it with real clinical practice. In this study both training and assessment took place in a highly controlled environment, i.e. the clinical skills laboratory. The real clinical workplace may, however, provide additional and different contextual challenges. The degree of training transfer from PFHS to real patients seen in this study therefore may not necessarily be replicable in the real clinical workplace. This is an important consideration as there remains ambiguity within the literature with respect to the degree to which simulation-based assessments conducted in the skills laboratory correlate to real clinical performance (Lynagh et al., 2007).

Finally, the findings in this study are derived from one training course and post-course assessment investigating one clinical skill, therefore further limiting the generalisability of findings. With respect to the course itself, the results need to be considered in line with its design and aim in that it is a short, focused course conducted in a single institution with the purpose of training basic cardiovascular examination skills and delivered with a high tutor to student ratio (2 to 6), to students who had no prior clinical examination experience. Equally, outcomes in terms of students' ability to perform a cardiovascular examination need to be interpreted in view of the assessment method, i.e. an OSCE circuit with limited stations involving SPs and patients with limited pathology.

6.5 Summary to Chapter

In this chapter, I presented the second main empirical study of the thesis comparing the use of PFHS to patients in clinical skills education, using the training and assessment of cardiovascular examination skills as a platform for inquiry. Within the limitations of the study, medical students with no prior cardiovascular examination experience and trained with PFHS (VS Group) in a short course performed similarly to those trained with real patients (RP Group) in a parallel short course, as well as 3rd year students (Clinical Group) who have had almost 1 academic year of ward-based experience. Students' performance in OSCE stations with VS/SP simulations directly correlated with performance in OSCE stations with real patients. This demonstrates evidence of concurrent validity of assessment of cardiovascular examination skills with VS/SP simulation. There was also a similar increase in self-rated confidence in performing cardiovascular examination on patients. Students value the usefulness of learning cardiovascular examinations skills with VS/SP simulations for foundation building, repetitive practice and integration of theory and practice. Key to the learning process was the role of the tutors in terms of providing role modelling of professional behaviour as well as instructional scaffolding. Whilst students perceived the VS/SP simulations to be useful, aspects of conducting the cardiovascular examination, particularly auscultation in VS/SP simulations were generally perceived to be easier than when performing on a real patient, which was primarily due to the Ventriloscope®. This may potentially reduce or create negative learning effects, though this potential effect was not observed in this study and further work will need to be conducted to clarify these issues.

CHAPTER 7 General Discussions

7.1 Introduction to Chapter

In this chapter I will provide an overall discussion of the findings of the thesis from the literature reviews and empirical studies with respect to the key research questions and explore their implications on the practice of SBE. To begin my discourse, I will first present an overview of key findings in the thesis so far.

7.2 Summary of Thesis Findings

In Chapter 1, I first presented the key principles of SBE in relation to healthcare education, including the nature of simulation, the drivers with respect to the current climate of medical education and the underpinning educational theories. Simulation offers in theory and in practice solutions to some of the major problems faced in current medical education. The ultimate aim of SBE is to improve clinical competence of healthcare professionals and ultimately result in better patient care. With the rapid adoption of SBE over the last two decades, there have also been increasingly creative approaches to simulating different clinical situations. PFHS was created in response to a number of issues identified in the practice of procedural skills simulation with PTTs in isolation, which only focuses on technical psychomotor elements of a clinical task and have been described as relatively "decontextualised", minimalist and inauthentic to real life practice. PFHS, a simulation technique which combines a SP and PTT to simulate "clinical encounters" authentically, aims to promote integrated performance, contextualised education, and emphasise on patient-centredness, thus re-establishing the link between simulation and clinical practice (Kneebone et al., 2006).

Despite the theoretical arguments for PFHS, its use in SBE as with any other simulation technique should be supported by empirical evidence. Therefore, in Chapter 2 I presented a focused systematic review of the current evidence of PFHS with respect to training effectiveness and assessment psychometrics. There is at present a limited empirical evidence base in relation to these areas. With respect to training effectiveness, positive learning effects for both technical and non-technical skills have been demonstrated, though inferences that can be drawn from the available studies are limited as learning effect was generally measured with subsequent simulations as opposed to workplace performance. For assessment, there is some evidence of construct validity. For example, a study has showed a positive correlation between performance in PFHS-based assessments to other validated assessments and clerkship performance (Isenberg et al., 2011). However, there was no evidence of concurrent or predictive validity of PFHS-based assessment found in this review. There is also a paucity of literature exploring the underlying processes of what happens during training and assessment with PFHS or simulation in general.

In Chapter 4 and 5, I investigated the value of PFHS as a means of introducing clinical context into simulation training of clinical skills using the clinical scenario of skin laceration management and closure on an A+E patient as a platform for my investigation. In Chapter 4, I investigated the use of context to modify clinical challenge in simulation-based assessment. In this study, I created two PFHS scenarios to reflect different levels of clinical complexity through an iterative process of cognitive task analysis. Novice clinician participants performed significantly worse than experts in the more complex scenario, but not the relatively simple scenario. Novices' performance was also significantly worse in the complex scenario when compared to the simple scenario, whereas expert clinicians' performance was not. These findings correlated to results from the post-simulation interviews, whereby participants' comments revealed that most novices found the complex scenario more challenging, whereas around half of the expert clinicians found it to be no more challenging than the simple scenario. In summary, the study demonstrated that, through using PFHS and systematic simulation design, the level of challenge of a given clinical task could be modified. This data adds to findings of a previous study which demonstrated that modification of context in PFHS scenarios can be used to subjectively increase the complexity (Higham et al., 2007), and provides some empirical evidence in support for use of the complexity risk matrix for simulation training proposed by Kneebone at al (2007) that I discussed in Chapter 1.

In Chapter 5, I investigated in more detail clinician participants' perception of assessment of their clinical competence with PFHS and how this compares to PTT-based simulations. I aimed to investigate finer details of their experience of the simulations in order to elucidate factors that may influence their behaviour and performance and perceptions of the simulations. Participants generally perceived the PFHS simulations to be realistic representations of real clinical practice and more authentic when compared to the PTT-based simulation. Although the PFHS scenarios were considered as realistic, there was demonstrable variation between the participants' perceived realism and artificiality in different elements of the simulation such as the wound prosthetic or the environment. These findings are in line with those found by Dieckmann et al (2007) who also identified relatively realistic and unrealistic elements to anaesthetic simulations, which they termed "reality" and "fiction" cues respectively. Despite these variably realistic/artifical components, patient interaction and the scenarios in which the clinical tasks were embedded were consistently perceived to be realistic. From the participants' perspective, it is also these which were central to triggering authentic behaviour and responses, with many stating that it allowed them to practice as they would in they usual clinical workplace. In addition, clinicians commented on the value of PFHS as an assessment tool to allow them to perform in an integrated manner, i.e. simultaneously exhibiting both technical and non-technical skills allowing them to demonstrate their ability to handle a "clinical situation", which the PTT-based scenario did not.

Despite the use of PFHS as an approach to increase authenticity of simulation-based assessments, a number of factors were identified that affected clinicians ability to perform and behave authentically. A number of these factors were due to the limitations in the design of the PFHS simulation study mentioned in the previous chapters. However, there were two key factors that hinder authentic behaviour that may be common to simulation-based assessments in general. The first issue relates to the effect of being assessed. A number of participants commented on how they behaved in a "simulatory" way. One participant described acting more "softly" than they might usually do in real clinical practice, even towards the obstructive patient in the complex scenario. These comments correlated with some observations where participants would shout out their thoughts or exaggerate movements which they may not necessarily do so in the normal clinical workplace. There appears to be two processes that are taking place here. One is that the participants, acknowledging that the purpose of the simulation is to assess their competence, want to demonstrate their ability thoroughly. Therefore, in addition to actions that are usually carried out in real world practice, assessees may also try to demonstrate areas that might not be as easily observable to show their underlying cognitive processes. This is often by exaggerating actions, which has been described to be akin to behaviour required to pass a driving test (Hunt et al., 2004). The second process is that the participants, as assessees, want to perform to meet the assessment criteria. In the case of the participant stating that they behaved more "softly", it may have been that they believed that the assessment requires them to conduct themselves in this manner in order to score highly according to the assessment criteria. Essentially, they are behaving in line with what they believe the examiner's idea of model professional behaviour is. Whilst this finding is one identified in this specific study on simulation-based assessments, it is an issue common to all assessments that require observation of performance, which I will discuss in more depth in section 7.4 (Schuwirth and van der Vleuten, 2011).

In Chapter 6, I presented the second main empirical study of this thesis investigating and comparing the use of PFHS to real patients for clinical skills training and assessment. I looked at the training and assessment of cardiovascular examination skills for medical students with PFHS simulations consisting of Ventriloscopes® and SPs (VS/SP) as a platform for my inquiry. In this study, pre-clinical students without previous experience of examining patients who were trained with VS/SP hybrid simulation (VS Group) performed similarly to those training with real volunteer patients (RP Group) in a post-training OSCE consisting of VS/SP and real patient assessment stations. In addition, they performed similarly to students who had gained almost a full academic year experience of clinical clerkship (Clinical Group). Further analyses comparing student's OSCE scores (checklist) of PFHS to real patient stations showed statistically significant direct correlation, i.e. concurrent validity. Interestingly, there is some evidence demonstrating the transfer of simulation training with patient manikin (e.g. Harvey) to performance on real patients (Butter et al., 2010; McKinney et al., 2013; Oddone et al., 1993). Therefore, although the study presented in Chapter 6 demonstrated transfer of skills from PFHS to a real patient, further research comparing learning outcomes of the existing types of simulations is also required to determine the best approach.

There were similarities in the learning process with PFHS and real patients in this study, both of which allowed not only deliberate repetitive practice, but also development of skills through instructional scaffolding under the tutors' guidance, which is a key consideration when designing simulation-training activities (Kneebone, 2005). The similarities between learning with PFHS and real patients are unsurprising given that they were designed to represent and match each other in the training course. Some differences were, however, also noted. As an educational tool, students commented on the usefulness of the Ventriloscope® as a standalone tool and some even elected to use it in isolation out of physical context during the course despite the availability of a simulated "patient" to practice on. On the contrary, the students in the RP Group were not able to engage in more focused auscultatory training out of clinical context during the training course.

Another key issue that students raised was that VS/SP simulations allowed for a greater error of margin when compared to performing the cardiovascular examination on a real patient. The effect of this is that students generally found examining easier in VS/SP simulations than on real patients. This could potentially result in potential effect of introducing negative learning as discussed. However, this did not translate into the findings of the OSCE assessment in that both RP and PFHS groups performed equally well. The reasons for this are not certain, but possible causes maybe that the particular OSCE assessment used in the study could not detect finer details of performance and was thus unable to discriminate between the groups of students. Alternatively even though the VP/SP simulations allowed a greater margin of error in performance, it may still have been sufficiently stringent to train students to the level of those trained with real patients. If the reasons are the later, then this may have implications on simulation design in terms of fidelity requirements for training, i.e. the degree of realism, which is an issue that remains inconclusive despite on-going research in this field (Aggarwal et al., 2010; Issenberg et al., 2011).

There were a number of limitations to both the systematic review and empirical studies, which were discussed in more detail within the respective chapters. Primarily, with respect to the empirical studies, most of the limitations were related to issues with sample size, and that studies were based on a single area of clinical competence, i.e. wound management and cardiovascular examination skills for the first and second studies, respectively.

In addition to issues sample size, the approach used to conduct the studies, i.e. mixed-methods also resulted in limitations to the qualitative arm of the studies. In pure qualitative research, data collection should ideally be continued to the point of saturation (Chapter 16, Cohen et al., 2007). However in the studies in this thesis, the number of participants interviewed was limited to those recruited for the quantitative arms of the studies. Even in Chapter 5, where all participants who took part in the simulation for the quantitative component of the study were interviewed, saturation could not be guaranteed.

A further limitation that needs to be considered is the background of the participants. Participants were predominantly students and doctors based in the UK who have undergone training and assessment in line with curriculums that may be quite different from those from other countries. For instance, most of the participants would be familiar with the format of OSCE-type assessments and have had some type of simulation-based training that is prevalent in current UK medical training(Howe et al., 2004). It is generally acknowledged that as well as clinical technique, good exam technique is just as crucial to performing well in an OSCE exam (Cooper et al., 2012). From my experience, there are also unspoken rules that can influence or dictate a participant's performance in a simulation, such as the need to engage in role play, suspend (or pretend to suspend) disbelief - in essence what some participants in the studies described as behaving in a "simulatory" way. This background-unspoken understanding of how to act, behave and perform in a simulation may have influence how a participant performs. Therefore, findings in this the studies may not be generalizable to other institutions and countries with different training systems.

With these limitations in mind, PFHS appears to have educational value in terms of producing a positive learning effect, and may have the potential to be comparable to training with real patients as seen in the second empirical study. From the perspective of assessment, it has been shown to have validity, correlating to performance on real patients. In addition, for clinical skills competence, it allows more holistic assessment and integrated performance. PFHS allows modification of context in which clinical skills are performed in order to alter the clinical challenge, which may add educational value by tailoring simulations to learners and assessors needs. By allowing challenge to be increased, it can also increase discriminating power and, as a result, validity (Kyaw Tun et al., 2012). Drawing on these studies and the literature, I will now discuss the wider implications of the findings and present theoretical and conceptual considerations on the role of PFHS and "contextualised" simulation in general on current SBE in healthcare.

7.3 PFHS, Contextualised Simulation, Training and Curriculum Considerations

From the findings of the empirical studies and the literature review, the evidence available so far in terms of learning effects and being able to simulate clinical challenge is encouraging and appears to support the use of PFHS for training, though there remain some important considerations. In the study in Chapter 6 when I compared PFHS and real patients for training of cardiovascular examination skills, the results demonstrate that training with PFHS achieves a similar level of competence to training with real patients. The results could equally be interpreted as that real patient training achieves a similar level of competence with PFHS. Another issue, which I raised earlier in section 6.5, is the finding that students, after exposure to the PFHS, chose to focus on practicing component tasks such as auscultation in isolation despite the availability of an SP on whom to practice their skills. A reason I suggested was that the students may have already gained sufficient understanding of the context and therefore began to concentrate on areas of the cardiovascular examination, which they found most difficult. In response to this, I then raised the question of whether it would be equally educationally beneficial to expose students to real patients (as opposed to PFHS) to allow them to gain an understanding of the context, followed by focused part-task simulation training of individual components. The reason I have raised these questions and issues is that, although in the literature and in my studies PFHS has been demonstrated to be effective for training, this does not immediately warrant its usage. Even if it is to be used, there remains the question of when and how.

In order to begin to answer these questions, one of the first issues that needs to be considered is the place of PFHS and contextualised simulation in the wider scheme of simulation training. Although there is at current little in the healthcare education literature that informs us of the use of relatively decontextualised and contextualised simulation, there is much work in the wider psychology and education literature on a similar issue, i.e. the use of part-task and whole task training for skills training. The potential value of this research on task training, on informing simulation training in healthcare has been recognised and drawn on by a number of author (Spruit et al., 2014). Whole-task training, whether it is in simulation or in workplace-based training, refers to training of a task or procedure in its entirety. In this respect, PFHS is essentially a form of whole-task training.

The evidence in the literature for when to use whole and part-task training is variable and not yet conclusive, but it is generally agreed to be task dependant. In the field of human factors, tasks are typically classified as discrete, continuous or serial (Schmidt and Lee, 1988). Discrete tasks are single units with a well-defined beginning and end, e.g. throwing a ball. Continuous are those with no set end or beginning, e.g. swimming and running. Serial tasks are those that contain a series of discrete components, such as performing a gymnastic sequence composed of a series of techniques. For discrete and continuous tasks, the current evidence favours whole-task training, whilst part-task training alone may be less effective or even give rise to negative learning. On the other hand, serial tasks can be divided into its respective components, i.e. part-tasks for more effective training (Lee, TD. et al in Karwowski, 2006). Another consideration as to whether learning is better achieved by part or whole-task training is the degree of dependence of the task components. If two components of a task are highly interdependent then the components should be practiced as a whole (Chapter 10, Spector et al., 2007). A final consideration lies with the complexity of a task. Studies have demonstrated that, for complex tasks, whole-task training yields better transfer of training in comparison to part-task (Chapter 20, Hancock et al., 2008). The reasons for this may be that performance of an overall complex task requires integration of its subcomponent tasks. Although whole-task has been shown to be more effective in complex tasks, part-task training is still advocated if there are aspects of the overall task that are too complex and may impair learning due to cognitive load. In this situation, a combination of part and whole-task training methods yields the optimum learning effect.

Drawing inferences from this research, for the training of a given clinical task, it can be seen that a mixed strategy is probably necessarily as it may consist of several task properties (discrete, continuous, serial, level of complexity), but that whole-task training is ultimately required for integration. For example, the wound scenario in Chapters 3 and 4 can be seen as a serial task that can be divided into sections, some of which are discrete such as injecting local anaesthetic and wound irrigation, whilst others are continuous, e.g. suturing. There is also likely to be a degree of interdependence between certain tasks, such as injecting local anaesthetic and patient communication. For PFHS, then, it may seem that its value lies in providing a means for whole-task training to allow integration of part-task components and maximise transfer of training. However, a major problem with trying to directly translate this research finding to simulation is that it focuses on a simulation as an individual educational entity without considerations of other co-existing learning environments and processes, in particular learning that takes place in the clinical workplace.

The need to strengthen the relationship between simulation training and clinical practice is recognised (Kneebone et al., 2004) . In recognition of the need to better align simulation and workplace training, Kneebone (2009), drawing on parallels with the training of performance artists and musicians, proposed a 3-staged model for surgical training, which he termed "Practice, Rehearsal and Performance". In this model, he emphasised the need for "contextualised simulation" (rehearsal) as an intermediary step of between the focused practice of component skills, such as technical skills on part-task trainers (practice), and practice within the clinical workplace (performance). Using small bowel anastomosis as an example, he proposed that surgical trainees should complete three complimentary, iterative and bidirectional stages of training, whereby 1) trainees initially train psychomotor aspects of bowel anastomosis to proficiency on part-task trainers (practice); followed by 2) training within a comprehensive simulated operating theatre environment (rehearsal), which may involve other team members and provides richer context; and finally, 3) training under supervision in the real operating theatre (performance). The reason for this proposed intermediary stage of "rehearsal" is the recognition that isolated training of component tasks do not adequately represent the contextual influences of real clinical practice. Based on this model, PFHS would fit into this intermediary rehearsal stage.

The benefits of the using PTTs to hone core psychomotor skills prior to practise on a real patient are obvious and supported by empirical evidence (Anastakis et al., 1999). However, one needs to question whether the proposed intermediary step of "rehearsal" with contextualised simulation is always necessary. Specifically, I argue that this model may be too simplistic in that it does not necessarily take into account the relationship between learning in the workplace and in simulation, and how learners' understanding of contextual influences is achieved. Taking the case of the small bowel anastomosis, a junior surgeon who is going to perform this procedure for the first time on a real patient should have had some general experience of being in a real operating theatre, whether as first surgeon or assistant, and ideally has had experience of at least in assisting in a small bowel anastomosis, and therefore should have some understanding of various contextual influences. In addition to this, there are likely to be a plethora of other educational influences taking place, such as guidance and supervision by an expert and the relatively controlled environment of a well-run operating theatre. For a junior surgeon who is already armed with this contextual knowledge and an understanding of how various task components need to be integrated in the real clinical world, as well as having close expert supervision available, the use of contextualised simulation within a comprehensive simulated operating theatre environment simply just to understand how it feels like to perform small bowel anastomosis in the context of an operating theatre environment may be superfluous. Furthermore, context may be introduced into training by less (physically) elaborate means, such as through mental imagery, a common technique used in other disciplines that require high levels of performance such as music and sport (Arora et al., 2011; Gregg and Clark, 2007).

There are however, certain types of clinical situations where a "rehearsal stage" may be beneficial. To illustrate this, I will draw on some work I had previously conducted, the literature and my own personal clinical experience. In a previous pilot project that was conducted together with several of my colleagues, we investigated the use of simulation for fire safety in the operating theatre in situ (Rollason et al., 2009). A key goal in the management of this situation is to ensure all staff and patients are safely evacuated. Despite having trained personnel involved, the evacuation of staff and patient proved challenging during the simulation. This is unsurprising given the presence of an anaesthetised patient, environmental constraints and a host of complex equipment, which provided numerous challenges. The findings highlighted areas where training, as well as policy, needed to be improved. On a similar theme, a study by Abrahamson et al (2006) looked at the use of simulation to identify issues with resuscitation protocols and training when applied to the context of a Severe Acute Respiratory Syndrome (SARS) cardiac arrest patient, which provides additional contextual challenges. A major contextual factor affecting performance (such as the time from cardiac arrest to initiating defibrillation) was that personnel involved in management of these patients were required to wear personal protective gear, which affected their ability to perform timely cardiopulmonary resuscitation.

Next, drawing from personal clinical experience as the on-call radiology registrar, I was once involved in a challenging cardiac arrest situation on a patient who was undergoing an MRI scan in the early hours of the morning. The first responders, myself and an Intensive Care Registrar had both previously attended generic simulation training for cardiopulmonary resuscitations (Advanced Life Support), as well as had real life experience in managing cardiac arrest situations. However, despite our combined experience, neither of us had any experience in managing a cardiac arrest in the MRI room, a very rare situation, and we both found the scenario highly challenging. The reasons for this is that the MRI room provides extra challenges to the process of cardiopulmonary resuscitation that may not be encountered in other clinical areas such as a ward. Vital equipment such as the cardiac arrest trolley and defibrillator cannot be brought into the MRI scanner due to MRI incompatibility. Therefore, we had to try and extract the patient out of the scanner, which inevitably took some time. During this time, we also had to think laterally how perform cardiopulmonary resuscitation without some of the usual equipment to minimise interruptions to the resuscitation process. In this example, a usually challenging task was made even more challenging due to the context in which it was performed. On personal reflection, I believe that if we had previously had not just generic simulation training in cardiopulmonary resuscitation, but training in the context of a patient arresting in an MRI scanner, then we may have been better prepared. Indeed, this training has been advocated by a number of authors (Gaca et al., 2007; Sica et al., 1999).

So, why may contextualised simulation be more beneficial in one situation when compared to the other? Looking specifically at two of the of clinical situations I described above, performing small bowel anastomosis under supervision in the relatively well-controlled environment of an operating theatre, and performing cardiopulmonary resuscitation on a patient in an MRI scanner, there are several fundamental differences. The first is how much of an understanding of the context of the working environment can one gain through daily practice in the workplace. Clearly, for the MRI cardiac arrest situation, there may be little scope to gain experience due to low clinical exposure. With respect to the small bowel anastomosis though, even if the junior surgeon had not previously performed the procedure, they may still have gained an understanding of contextual influences through observation as described in Bandura's social learning theory (Bandura, 1976). Second, is the presence of external factors that can compensate for relative lack of understanding of contextual influences, such as the support a trainee gains from a supervisor when they perform a task for the first time on a real patient. Indeed, there is empirical evidence demonstrating that trainee surgeons under expert supervision can perform operations safely with equivalent patient outcomes to those of fully trained surgeons (Acun et al., 2004; Crolla et al., 1997).

In this discourse on the role of PFHS and contextualised simulation in training, I have tried to raise some key issues in the current research of simulation-based training, issues that also need to be considered when interpreting the results of the empirical studies in this thesis. Much of current research on the effectiveness of training looks at simulation as a separate entity to other forms of training. The reasons for this are likely to be related to constraints in research design, as well a strong emphasis on controlling for variables. For instance, there are some studies that employ medical students or absolute novices to investigate learning curves for procedural skills training with simulation, the reasons, though not always stated explicitly as such, appear to be controlling for prior clinical exposure and experience (ref(Aggarwal et al., 2006; Eversbusch and Grantcharov, 2004). This is by no means a criticism on the research methods, but these research approaches limit the generalisability and transferability of findings to the rather more *messy* world of medical education. Also, as mentioned earlier, proven effectiveness in a study does not necessarily warrant usage of an educational intervention. Just because a clinical situation can be successfully simulated, it doesn't mean it should be. Simulation should be used when it provides the most educational utility and value as measured in relation to other coexisting educational activities.

Returning to the question I asked earlier at the beginning of this section of whether we should just expose students to real patients and then allow them to practice components skills on PTT-based simulations as opposed to PFHS, the short answer is that it depends. For example, in the medical school where I studied, there are over 300 students per year and clinical exposure is often variable. The clinical workplace may also give rise to constraints in learning. Taking these into account, the use of PFHS is probably justified in this situation. In sum, in order to understand the role of PFHS and contextualised simulation in current medical education, further research and theoretical development is required to allow us to place them in a wider, more comprehensive map of educational activities.

7.4 PFHS and Assessment Considerations

From the empirical studies in this thesis and within the current literature, there is some evidence that demonstrates PFHS as a valid and reliable approach to assessing clinical competence. However, as with the discussion in the last section regarding suitability of PFHS for training, these findings do not automatically warrant its use in assessment of clinical competence. In order to determine its role in assessment of competence a key factor that needs to be considered is its educational utility. A commonly described model for measuring assessment utility in medical education is described by Van de Vleuten (1996) as follows:

Utility = educational impact x reliability x validity x feasibility x acceptability

Using this model, it can be seen that utility is dependant on a number of contributing factors. Although investigating the utility of PFHS was not a key research question in this thesis, some of the findings in this thesis do provide support for its use in terms of providing evidence of validity, reliability, and acceptability. However, educational impact and feasibility remain to be determined. Of note, PFHS inherently has higher costs in comparison to more conventional PTT-based simulations and the question of feasibility was raised by a number of clinicians interviewed in Chapter 5 regarding the skin laceration management PFHS scenarios. At the same time, PFHS is probably more feasible for assessing clinical competence for some clinical scenarios, particularly those that are rare or involve invasive elements in comparison to workplace-based assessments.

With respect to educational impact, as mentioned in Chapter 1, it is widely acknowledged that assessment has the potential to drive learning. Learners tend to learn what is set in an assessment criteria to pass an assessment (Wormald et al., 2009). Feedback from an assessment can also provide learners with insight into their strengths and weaknesses. However, aside from the positive effects on learning, it has also been demonstrated that assessments can give rise to negative and unintentional learning effects. For example, a study by Rudland et al (2008) investigating the effect of OSCE on learning strategies demonstrated that students elected to concentrate on studying checklists and practicing on their flatmates, rather than increasing workplace based training. A key reason for this was that the OSCE stations in this specific study were noted to be inauthentic to real clinical practice and instead encouraged students to work on exam techniques such as being able to complete a simulated clinical task comprehensively within the time constraints of an OSCE examination. The clinical workplace was therefore not perceived by students to be the most efficient environment to prepare for OSCEs. These findings are however not entirely negative. The OSCE still exerted a strong effect of motivating students to learn. However, the fact that it draws students away from obtaining real clinical experience is worrying in that it may not encourage students to develop some of the less explicit (i.e. within the “hidden curriculum”), but equally important skills for real professional practice. Taking these into consideration, in theory, PFHS and contextualised simulation therefore has the potential to drive learners to develop their clinical competence by being authentic to clinical practice, though the true effect of this has yet to be demonstrated empirically.

Another key property of PFHS, which does not come into the above equation for assessment utility, is its ability to increase authenticity. As demonstrated in Chapter 5, clinicians generally perceived PFHS-based assessments to be authentic to real clinical practice, though there remains the question of what value this may have. A possible answer is its role in assessment validity. To understand this, we need to first discuss the nature of validity and authenticity and their relationship. Many proponents of the authentic assessment paradigm argue that, by nature of approximating real life practice, it infers strong validity (Archbald and Newmann, 1988; Wiggins, 1993). However, the terms validity and authenticity are not synonymous and should not be confused with one another. Validity relates to whether the assessment measures what it intends to measure. It is therefore a relative measure in relation to a defined assessment criterion. For example, as discussed in Chapter 2, taking the assessment of competence of suturing skills, if the assessment purpose and criteria is to assess basic psychomotor skills, then assessing someone's ability to suture a banana may be equally valid to suturing a highly realistic skin simulator (Kyaw Tun et al., 2011). Authenticity of an assessment, on the other hand, is referenced to real life clinical practice. If the criteria of an assessment are referenced to real life clinical practice, then, by increasing authenticity of a simulation-based assessment, validity should also be increased. In this respect, the use of PFHS should be dependant on the purpose of the assessment, which, if is to assess real life clinical competence, then may be beneficial in terms of validity. This is supported to an extent by some of the study results in this thesis. In Chapter 6, the PFHS were designed to match the real patients in terms of signs, symptoms and patient interaction, which may have resulted in the strong correlation in the performance ratings seen between the two types of OSCE stations.

There may be further value of increasing authenticity beyond content validity, i.e. comprehensiveness of an assessment of the different aspects of a clinical competence for a given clinical skill. If the sole purpose of the assessment was to ensure that there was adequate sampling of the individual component skills required for clinical competence, one could argue that a series of more focused assessments to measure each component individually should suffice. However, the combined simultaneous assessment of the various components of clinical competence may produce effects that are different than when just assessing each component skill alone in series. This is an area that requires further empirical research.

There is a finding in this thesis, which may support the notion that there is value of authenticity beyond increasing validity. In the study in Chapter 5, clinicians stated that in the PFHS scenarios, by allowing them to simultaneously exhibit various component skills as they would in real clinical practice, the simulation allowed them to behave more in line with how they would do in usual clinical practice. The issue of trying to capture authentic clinical behaviour in assessment may be an important one, particularly if trying to determine aspects of performance such as professionalism. It is also worth noting that the issue of measuring professionalism is also another area where research is lacking, perhaps in part due to the inherent difficulty in trying to define and assess this aspect of performance (Riley and Kumar, 2012).

Of note, some clinicians commented on the difficulty in performing well or producing errors when performing in the PTT-based simulation due to its relatively inauthentic nature. This issue of the inauthenticity of some simulations giving rise to potentially negative performance has been raised previously (Fidment, 2012). This inauthenticity of the simulation-based assessment may require assesses to adopt strategies or “exam technique” not directly related to clinical competence in order to do well. This is potentially a wider issue not limited to simulation but assessments in general. From my observations as a teacher and examiner, it is not uncommon for assessments to require assesses to be able to exhibit a degree of “exam technique” in order to pass, but this should not be the emphasis of the test. Instead, an assessment should ideally be designed to account for or minimise requirements for exam technique, such that assesses are free to demonstrate their true ability. In this respect, increasing authenticity to allow clinicians to behave as they may do in real clinical practice may be of benefit.

Finally, as discussed in Chapter 1, the authenticity of an assessment should provide value beyond the classroom, which in the case of healthcare education is lifelong professional clinical practice. This requires an assessment to not only reflect a real life activity, but also promote and foster desired values that beyond the actual activity itself. “Contextualising” simulation such as with this technique of PFHS, which appears to bring about more meaningfulness, allowing users to relate to their professional identity, may be one way of achieving this. In this respect, it is aligned with some of the principles of authentic education discussed in Chapter 1.

Taking the issues of utility and educational impact of assessment discussed above into consideration, there is a general consensus that no single assessment method is always superior to another, and that a multimodal, multidimensional approach is required (Epstein and Hundert, 2002). The use of PFHS, as with other forms of assessment, should be aligned with the desired learning and curriculum objectives. Furthermore, to understand the role of

PFHS-based assessment, it needs to be considered in light not only of its own individual utility, but also in relation to other available assessment methods currently used.

7.5 Proposal for Uses of PFHS and PTT

Have considered the implications of the research presented in this thesis, I will now present a proposal for use of PFHS and PTT in current healthcare education based on the discussions so far. Given the limited evidence in this field, these proposals aim to be at most suggested guidance on usage and are by no means prescriptive, but should be applied with a degree of pragmatism. The proposals presented are therefore as much for suggested use, as for a platform for further scholarly debate.

7.5.1 Suggestions for use of PFHS and PTT in Training

- PTTs should be used when the purpose of training is to improve component skills. This should not be only for absolute novices, but also for advanced professionals when they need to focus on component skills as part of "deliberate practice". For example, like with a concert level musician who may still need to practice difficult parts of their pieces (i.e. component tasks), a qualified surgeon may still need to practice and maintain difficult component skills such as laparoscopic suturing.
- PFHS should be used as a platform for trainees to practice component skills in an integrated manner to better mimic clinical practice, but only when those component skills have been acquired.
- The decision to use PFHS should be made in light of other co-existing educational activities, such as learning in the workplace and where an understanding the influence of context and the challenges it brings on a given procedure can be acquired effectively and safely. The use of PFHS should be especially considered when training in the workplace alone is limited, impossible and/or unsafe to staff and patients.
- To ensure maximum educational benefit, the use of both PTT and PFHS should be carefully mapped onto a curriculum. The curriculum itself should strongly mirror the needs of real world professional practice.

7.5.2 Suggestions for use of PFHS and PTT in Assessment

- The use of both PTT and PFHS should be linked to desired assessment criteria and not merely be used because the technique is available.
- When the purpose of assessment is to evaluate component skills only, then PTT simulations may be used. However, when the decision is made to assess component skills only, efforts should be made to ensure it remains meaningful and strongly aligned to the assessee's professional practice.
- The use of PFHS should be considered when a) the assessment of integrated performance is required and b) when there is a need to introduce contextual challenges.
- PFHS should be considered especially when there is a need for holistic assessment of competence of a given procedure within a controlled environment. This may be for the purpose of a high stakes examination when reproducibility is a key consideration. It may also be for clinical procedures or scenarios, where alternatives such as workplace-based assessment are impractical or impossible.
- The use of PFHS for assessment is most likely to of benefit when used as part of a mixed strategy incorporating a variety of tools for assessment of competence. For example, within an OSCE exam circuit, a combination of stations based on PFHS and PTT simulations may be considered such that the benefits from both types of simulation are taken advantage of, i.e. to produce a comprehensive picture of an assessee's overall competence when used in combination.
- In addition to use in high stakes examinations, the use of PFHS should be considered in formative assessment, combined with debriefing and feedback to help students identify their learning needs. Scenarios that are more challenging than what a trainee may be expected to encounter in their usual practice may also be considered for training purposes.
- Given the positive but limited evidence for the use of PFHS, if and when introduced into high stakes examinations, its use should be subject to rigorous evaluation to ensure the desired educational outcome.

7.6 Personal Reflections

In this section, I am going to present a series of more personal reflections and opinions drawn from my experience of developing this thesis, and my involvement in SBE in general. Researching SBE is challenging. This is clear, as despite the exponential increase in the available research, the quality of studies is often variable, usually due to study constraints and there remains a demand for more and higher quality research (Bradley, 2006; Cook, 2010). A major contributing factor to the difficulty of researching SBE is that fundamentally, much of the research into SBE requires measurement of multiple aspects of human performance. Unlike studies designed to tests new drugs, where trials are relatively well controlled, researching human performance and behaviour which is far less controllable is much more challenging (Chapter 1, Martin and Bateson, 1993). This is in part due to the wide variation in performance and definition of performance standards. Whilst psychomotor elements of performance, e.g. suturing ability, may be relatively easy to measure, domains such as professionalism and good communication are not yet clearly defined, and may never be (Riley and Kumar, 2012). Although these issues were often a cause of frustration when conducting my research, it is also these issues that make that research in SBE is particularly interesting. Through these reflections, I hope to give my readers insight into some of the practical issues of conducting studies in SBE, limitations and their underlying reasons, with a view to informing better research practice.

7.6.1 Introducing Inauthenticity into Authentic Assessment for Research

Whilst conducting my empirical studies, I found that there were tensions between trying to create simulations that were authentic to clinical practice, but at the same time allowed for comprehensive assessment of my participants' competence. I first encountered this issue when piloting the wound management PFHS scenarios for the study in Chapters 4 and 5. One of the relatively novice participants whom I perceived to be at least at his level of expected competence, when faced with the complex PFHS scenario (drunken uncooperative patient), maintained a high level of professionalism attempting to develop rapport and patient co-operation. However, when attempting to perform wound closure, he was continually faced with repeated verbal abuse from the uncooperative SP and decided that it was unsafe to proceed in performing an invasive medical procedure (suturing). As a result, the participant explained in a professional manner to the (simulated) patient that he would perform the procedure when the patient was more sober and would also ask a more senior colleague to help for reasons of patient safety. The participant effectively prematurely terminated the simulation before he could demonstrate the range of skills of interest necessary for the study. From my clinical experience, this is an entirely valid and safe approach to the clinical scenario that was presented to him. I would even commend him on his professionalism if I saw this during his actual clinical practice. However, this was clearly problematic for the study, as the scenario was not “completed”.

This raises two issues in terms of the design of simulation-based assessment. First, if the purpose of simulation-based assessment is to measure a well-defined set of competencies, then the authenticity of simulation may need to be compromised by introducing constraints in the design such that participants will (or will attempt to) demonstrate those competencies. In the case of this particular study, the participant’s briefing was changed so that it was clearer and more specific as to what the assessment would like them to achieve, such as wound closure. On the contrary, if we want to conduct simulation-based assessments which allow participants to have a greater degree of freedom in terms of how they approach a clinical problem, i.e. allowing for different, but equally valid solutions, the assessment criterion and performance metrics must themselves be accommodating and flexible enough. These, of course have to be balanced with the goals of assessment in that, if there are very specific competencies that need to be assessed, flexibility of the assessment criteria may need to be sacrificed.

7.6.2 Assessing performance or performance for assessment?

Earlier in Chapter 5, I discussed an issue with the PFHS assessments in that, despite being designed to increase authenticity, participants still exhibited what one of them described as “simulatory” behaviour, such as trying to demonstrate more empathy towards a patient than they normally would do. I believe the main reason for this behaviour is largely due to participants’ knowledge of being observed and assessed. Although I raise this issue in relation to simulation-based assessments, it is actually an issue with all assessments where a participant is aware that he/ she is being observed, such as in workplace-based assessments (Schuwirth and van der Vleuten, 2011). For example, in a study examining workplace based assessments for anaesthetists, Castanelli and Kitti (2011) found that it was often difficult to observe suboptimal behaviour as the participants ensured they would “be on their best behaviour” when they knew they were being assessed. This is problematic in that it inhibits the capturing of authentic behaviour of students and clinicians. There have been attempts to minimise the observer effect. Video recording, which removes the physical presence of an assessor, has been suggested to reduce the observer effect (Kneebone, 2006). However, as can be seen in the studies in this thesis that the observer effect appears to still very much present.

An approach to this problem is to use of unannounced incognito SPs, which has been used to assess general practitioners in the primary care clinic setting. Here, the general practitioners are unaware that the patient is in fact an SP as he/she portrays a patient accurately (Gorter et al., 2002). Of course, this technique is limited from a practical perspective in that intimate and invasive procedures cannot be simulated. My reasons for raising this issue are not so much to dismiss the usefulness of simulation-based assessments, but to highlight one of the major barriers to assessing and researching behaviour and professionalism with this method. In light of these barriers, the inferences drawn from current studies assessing behaviour and performance with simulation need to be interpreted with care.

7.6.3 Limitations to observing competence

One of the problems I encountered during the studies which I have also observed in my practice as an OSCE examiner, is that, observation of performance in the simulation alone is limited in its ability to actually determine one’s competence. As a regular OSCE examiner, at times I have been instructed to just let the students perform the necessary examination without opportunity to question or probe more deeply the students’ underlying thoughts and cognitive processes. The reasons for this are to maintain objectivity and standardisation of assessment. Equally in the studies presented in this thesis, performance of students and clinicians were recorded on video with no scope for questioning them. However, this raises some problems. Aside from not being able to probe an assessee’s underlying cognitive processes, assessees may outwardly seem to know what they are doing to the assessor, but may in fact not. For example, as mentioned in Chapter 6 during one of the focus groups, an interesting dialogue was captured whereby students stated that during the OSCE assessment of the cardiovascular examination skills, they performed the action of "looking" for various signs of cardiovascular disease in the patient's hand and even shouted that they were doing so, but in fact admitted to not knowing what those signs actually look like. In this case, the patient actually did not have any signs of cardiovascular disease in the hands, and the students stated that there were none. However, the examiner seeing these actions may believe that the students actually knew what they were looking for. This draws me to another major consideration when interpreting results of research in simulation. The performance of study participants measured within the studies, even with the use of validated rating tools, may not necessarily reflect their true competence.

7.6.4 Validity and transfer of training to clinical practice - in search for the Holy Grail?

Two of the major challenges in researching SBE and ones that is recurrently mentioned in the literature are 1) how to determine predictive validity for simulation-based assessments, i.e. direct comparison to real clinical performance, and 2) how to determine transfer of training from simulation to real clinical practice. For assessment validation, ideally, performance in simulation and clinical practice should be measured and compared for predictive or concurrent validity. For transfer of training, it has been suggested that randomised controlled trials, comparing effect of simulation training to no-simulation training on performance in real clinical practice is required (Lynagh et al., 2007; Sutherland et al., 2006).

However, this approach to studying simulations raises a number issues. From a practical perspective, if the clinical procedure is rare, then it may be difficult to evaluate transfer of training or to assess performance in the clinical workplace for validation purposes. If there is a long time before the simulation and workplace-based assessment, then a number of confounding factors may arise, such as an interval increase in competency between the two assessments.

There are also ethical issues of concern. If someone performs badly in a simulation, would it be ethical to allow him or her to still perform on a real patient for the purpose of validation? Also, would it still be ethical to allow a randomised group of students or clinicians who had not received simulation training to practice on a real patient for the purpose of research? During a conference I attended some years ago, I raised this issue to a number of my peers, some of whom replied that it is ethical to conduct randomised controlled trials comparing no simulation training to simulation training for invasive procedures on real clinical practice, as there is “no baseline evidence of its efficacy”, adopting the same ethical stance as they would for a drug trial. However, I contest this argument primarily as in addition to existing strong theoretical evidence for the use of simulation, there is a small, but steadily growing body of empirical evidence that demonstrates a degree of concurrent or predictive assessment validity, and transfer of training (Cook DA, 2011). I believe that the stance “we do not know that it works so we can randomise” is wrong and should not be adopted when patient harm may possibly result.

Instead, to overcome this ethical dilemma, there are approaches to studying simulation that may give us some of the empirical evidence needed without potentially jeopardising patient safety and care. One approach that has been used in the simulation literature is to review performance of cohorts of trainees and students before and after an educational intervention. For example, Barsuk et al (2009) evaluated CVC line complications rates in a period before and after introducing simulation training and demonstrated a change towards better practice behaviour (use of ultrasound) and reduced complications. Of course this approach has its own limitations in that 1) it is retrospective; 2) cohorts are studied in succession rather than in parallel; and 3) cohorts may not be matched, all of which can potential give rise to confounding factors.

The other approach, which I employed in this thesis, is to conduct controlled trials comparing simulation to real patient for training and assessment of clinical procedures that are unlikely to cause patients harm. In this thesis, I specifically used the training of cardiovascular examination skills, as it is non-invasive. From a practical perspective, selecting to study cardiovascular examinations skills is also quite feasible, as patients with the required cardiovascular signs were relatively common and accessible. Although much of the drive for using simulation is focused on training for rare events and invasive procedures, I believe that a systematic research programme that studies a range of relatively common non-invasive clinical procedures may provide us with the generalisable theories and models required to inform educators of not only if, but also how and when to use simulation in the wider field of healthcare education.

With respect to justifying use of simulation, there is also a demand for every new simulation design to be “validated” (Schout et al., 2010). This is also problematic, as it requires a retrospective evaluation of the simulation, often after a lengthy design process, particularly when it comes to simulations using sophisticated technology. Current validation methods are also limited in that they do little to inform what aspects of simulation design makes it valid. For example, a common approach for construct validation is through differentiation of experts and novices on the theoretical basis that experts perform better than novices (Carter et al., 2005; Schout et al., 2010). There are two issues with this. First, a simulation that does not meet all the necessary content of the assessment criteria may still differentiate relative novices and experts. Second, there are issues with expert and novice classifications. Some studies compare consultant level clinicians to medical students, which is hardly a fair comparison (Arora et al., 2005; Kenney et al., 2009; McDougall et al., 2006; Woodrum et al., 2006). In addition, many of these studies use experience level as a means of classifying experts. However, particularly for competency domains in communication and professionalism, this assumption may not be true (De Vries et al., 2014; Kim and Myung, 2014). Anecdotally from my experience, many healthcare professionals including myself have experience in interacting with senior clinicians who may not exhibit the highest level of professionalism.

Finally, in addition to researching transfer of training and validity, another area that is currently lacking in the field of SBE research is studies investigating how simulation works, not just macroscopically (for e.g. how it promotes deliberate repetitive practice), but also microscopically, looking at the finer details of underlying mechanisms and causal effects. This is something I attempted to do to in this thesis. For example in Chapter 5, with respect to PFHS, although the simulations were generally designed to be authentic to clinical practice, there were still a host of artificial elements. Overall though, these weren’t always important to the clinicians. However, it appeared that the crucial element to PFHS that promoted authentic behaviour was the presence of a real human being simulating patient interaction allowing participants to approach the simulation as they would in real clinical practice.

7.6.5 A tool is only as good as…

In systematically researching PFHS during this thesis, my findings have demonstrated that it has educational value for training and assessment. However, the findings in this study must be interpreted in light of the context of the study and simulation designs. Earlier in the Chapter, I discussed how the educational value of PFHS could only be determined by considering its place in the wider context of healthcare education. In addition to this, PFHS as with other approaches to simulation is only a technique and its usefulness is dependant on how it is employed. I have seen variable approaches to PFHS in current clinical skills training, some of which may not necessarily give rise to best educational practice.

For example, there are times when the SP may not be highly trained in their craft, or where the SP is replaced by a member of teaching faculty or students, usually for practical or cost reasons. Sometimes there is little thought put into the development of SP roles and clinical scenarios. Although in these examples, there may still be some educational value, in my opinion, it is perhaps not the best way to use PFHS. My reasons are that the underlying principle and greatest strength of PFHS is to create authenticity in particular with relation to patient interaction. Yet, the examples I gave above, risk trivialising this important property of PFHS. In this thesis, where possible, I used professional SPs who had experience in portraying real patients and spent a significant amount of time designing scenarios to maintaining authenticity. In this respect, my findings may not apply to when PFHS is used in what can perhaps be considered to be a suboptimal manner.

7.7 Conclusion

PFHS, which combines SPs and PTT is a relatively novel approach to simulation designed to address some issues in current SBE and was introduced in response to issues raised with more conventional approaches to simulation of procedural skills. These issues were primarily related to the relatively decontextualised nature of isolated component task training and assessment, which did not take into account the complexities of real clinical practice. The conceptual and theoretical arguments underpinning PFHS are that it allows integrated practice and performance of component skills and allowing tasks that have previously been simulated in isolation to be performed in context. However, early in this thesis, I conducted a focused systematic review of the literature, which demonstrated positive, but limited evidence to support its use. This in turn provided the basis for my empirical inquiry into the use of PFHS as an approach to simulation-based training and assessment of clinical skills.

The findings within this thesis provide further evidence in addition to the existing empirical work in the literature supporting the use of PFHS due to its properties in terms of increasing authenticity and context. From the perspective of assessment, a lack of authenticity may not only diminish validity (if the assessment criteria is referenced to real clinical performance), but may also cause unintentional effects to performance, e.g. inducing errors due to the assessment methods. The context and authenticity provided by PFHS also has value in terms of setting assessment of clinical skills in a more naturalistic framework. Crucially, the value of a real human being within the scenario appears to be beneficial from the perspective of simulating patient interaction authentically and allowing clinicians to "connect" with the simulation in terms of behaviour and performance. Another key benefit demonstrated in this thesis is the ability to modify clinical context to allow different levels of clinical challenge. From a training perspective, it is demonstrated within this thesis that it may allow students to achieve similar levels of basic competency in clinical skills to when training with patients, and therefore has the potential to augment clinical training.

The implications of these findings are, however, limited. This is partly due to the individual study limitations, but also some of the problems with respect to research in SBE in general that I highlighted, in the previous section. Reflecting on these findings and considerations, there is much scope for further research into the use of PFHS. First, there needs to be more work investigating the use of PFHS for training and assessment across different clinical skills to provide an empirical basis for developing more generalisable theories. Initial directions for such research should include studies exploring the use of PFHS not only across a wider range of clinical skills, but with healthcare professionals and students across a wider range of disciplines. In addition to this, more research and theoretical development is needed to understand how and when PFHS should be used in the wider context of healthcare education in general, and particularly in relation to workplace education, such that it provides maximal educational benefit. For training, PFHS and contextualised simulation is likely to be most beneficial where there are barriers to effective learning in the workplace, such as when there is low clinical exposure, or when constraints in the workplace limit skills acquisition. For assessment, the use of PFHS must be used in line with not only the goals of assessment, but in consideration with overall utility. Finally, I argue that more research is needed to understand the nature of the practice of simulation, rather than just outcomes. For example, as demonstrated in Chapter 6, on face value, simulation provides a means for skills acquisition through deliberate repetitive practice, but a closer look demonstrated that, much of the learning was enhanced by instructional scaffolding and tutor role modelling. Likewise, in Chapter 4 and 5, I set out to design simulations with a high degree of authenticity for assessment of competence, but in doing so also demonstrated constraints brought about by increasing authenticity. To address these research needs, a diverse research strategy should be employed in order to better determine the rational use and design of PFHS.

References:

Abrahamson, S., Canzian, S., Brunet, F., 2006. Using simulation for training and to change protocol during the outbreak of severe acute respiratory syndrome. Crit. Care 10, R3. doi:10.1186/cc3916

Acun, Z., Cihan, A., Ulukent, S.C., Comert, M., Ucan, B., Cakmak, G.K., Cesur, A., 2004. A Randomized Prospective Study of Complications Between General Surgery Residents and Attending Surgeons in Near-Total Thyroidectomies. Surg. Today 34, 997–1001. doi:10.1007/s00595-004-2857-7

Aggarwal, R., Grantcharov, T., Moorthy, K., Hance, J., Darzi, A., 2006. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. Am. J. Surg. 191, 128–133.

Aggarwal, R., Mytton, O.T., Derbrew, M., Hananel, D., Heydenburg, M., Issenberg, B., MacAulay, C., Mancini, M.E., Morimoto, T., Soper, N., Ziv, A., Reznick, R., 2010. Training and simulation for patient safety. Qual. Saf. Health Care 19, i34–i43. doi:10.1136/qshc.2009.038562

Alam, U., Asghar, O., Khan, S.Q., Hayat, S., Malik, R.A., 2010. Cardiac auscultation: an essential clinical skill in decline. Br J Cardiol 17, 8–10.

Alessi, S.M., 1988. Fidelity in the design of instructional simulations. J. Comput.-Based Instr.

Alinier, G., 2007. A typology of educationally focused medical simulation tools. Med. Teach. 29, e243–250. doi:10.1080/01421590701551185

Anastakis, D.J., Regehr, G., Reznick, R.K., Cusimano, M., Murnaghan, J., Brown, M., Hutchison, C., 1999. Assessment of technical skills transfer from the bench training model to the human model. Am. J. Surg. 177, 167–170. doi:10.1016/S0002-9610(98)00327-4

Anders Ericsson, K., 2008. Deliberate Practice and Acquisition of Expert Performance: A General Overview. Acad. Emerg. Med. 15, 988–994. doi:10.1111/j.1553-2712.2008.00227.x

Anderson, J.R., Reder, L.M., Simon, H.A., 1996. Situated learning and education. Educ. Res. 25, 5.

Ansell, J., Mason, J., Warren, N., Donnelly, P., Hawkes, N., Dolwani, S., Torkington, J., 2012. Systematic review of validity testing in colonoscopy simulation. Surg. Endosc. 1–13.

Archbald, D.A., Newmann, F.M., 1988. Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School.

Arora, H., Uribe, J., Ralph, W., Zeltsan, M., Cuellar, H., Gallagher, A., Fried, M.P., 2005. Assessment of construct validity of the endoscopic sinus surgery simulator. Arch. Otolaryngol. Neck Surg. 131, 217–221.

Arora, S., Aggarwal, R., Sirimanna, P., Moran, A., Grantcharov, T., Kneebone, R., Sevdalis, N., Darzi, A., 2011. Mental practice enhances surgical technical skills: a randomized controlled study. Ann. Surg. 253, 265–270.

Asghar, O., Alam, U., Khan, S., Hayat, S., Malik, R.A., 2010. Cardiac auscultation: the past, present and future. Br. J. Cardiol. 17, 283.

Bandura, A., 1976. Social Learning Theory, 1st edition. ed. Prentice-Hall.

Barsuk, J.H., Cohen, E.R., Feinglass, J., McGaghie, W.C., Wayne, D.B., 2009. Use of simulation-based education to reduce catheter-related bloodstream infections. Arch. Intern. Med. 169, 1420–1423.

Barsuk, J.H., McGaghie, W.C., Cohen, E.R., O'Leary, K.J., Wayne, D.B., 2009. Simulation-based mastery learning reduces complications during central venous catheter insertion in a medical intensive care unit*. Crit. Care Med. 37, 2697.

Batchelder, A.J., McCarthy, M.J., 2013. a critical appraisal of the evidence of current methods of assessment for surgical trainees. Bull. R. Coll. Surg. Engl. 95, 7–11.

Bezemer, J., Korkiakangas, T., Weldon, S.-M., Kress, G., Kneebone, R., 2016. Unsettled teamwork: communication and learning in the operating theatres of an urban hospital. J. Adv. Nurs. 72, 361–372.

Black, N., 1994. Why we need qualitative research. Br. Med. J. 48, 425.

Black, P., Wiliam, D., 1998. Inside the black box: Raising standards through classroom assessment. Granada Learning.

Bond, W.F., Lammers, R.L., Spillane, L.L., Smith-Coggins, R., Fernandez, R., Reznek, M.A., Vozenilek, J.A., Gordon, J.A., 2007. The use of simulation in emergency medicine: a research agenda. Acad. Emerg. Med. 14, 353–363.

Boursicot, K., Roberts, T., 2005. How to set up an OSCE. Clin. Teach. 2, 16–20. doi:10.1111/j.1743-498X.2005.00053.x

Boyle, E., O'Keeffe, D.A., Naughton, P.A., Hill, A.D.K., McDonnell, C.O., Moneley, D., 2011. The importance of expert feedback during endovascular simulator training. J. Vasc. Surg. Off. Publ. Soc. Vasc. Surg. Int. Soc. Cardiovasc. Surg. North Am. Chapter 54, 240–248.e1. doi:10.1016/j.jvs.2011.01.058

Bradley, P., 2006. The history of simulation in medical education and possible future directions. Med. Educ. 40, 254–62. doi:10.1111/j.1365-2929.2006.02394.x

Bradley, P., Postlethwaite, K., 2003a. Simulation in clinical learning. Med. Educ. 37, 1–5.

Bradley, P., Postlethwaite, K., 2003b. Setting up a clinical skills learning facility. Med. Educ. 37, 6–13.

Bray, B.S., Schwartz, C.R., Odegard, P.S., Hammer, D.P., Seybert, A.L., 2011. Assessment of human patient simulation-based learning. Am. J. Pharm. Educ. 75.

Butter, J., McGaghie, W.C., Cohen, E.R., Kaye, M., Wayne, D.B., 2010. Simulation-based mastery learning improves cardiac auscultation skills in medical students. J. Gen. Intern. Med. 25, 780–785. doi:10.1007/s11606-010-1309-x

Carter, F.J., Schijven, M.P., Aggarwal, R., Grantcharov, T., Francis, N.K., Hanna, G.B., Jakimowicz, J.J., 2005. Consensus guidelines for validation of virtual reality surgical simulators. Surg. Endosc. Interv. Tech. 19, 1523–1532.

Castanelli, D., Kitto, S., 2011. Perceptions, attitudes, and beliefs of staff anaesthetists related to multi-source feedback used for their performance appraisal. Br. J. Anaesth. 107, 372–377. doi:10.1093/bja/aer152

Ceci, S.J., Roazzi, A., 1994. The effects of context on cognition: Postcards from Brazil. Mind Context Interactionist Perspect. Hum. Intell. 12, 74–101.

Chikwe, J., Souza, A.C. de, Pepper, J.R., 2004. No time to train the surgeons: More and more reforms result in less and less time for training. BMJ 328, 418–419. doi:10.1136/bmj.328.7437.418

Clark, J.M., Maben, J., Jones, K., 1997. Project 2000. Perceptions of the philosophy and practice of nursing: preparation for practice. J. Adv. Nurs. 26, 246–256.

Cleland, J.A., Abe, K., Rethans, J.-J., 2009. The use of simulated patients in medical education: AMEE Guide No 42 1. Med. Teach. 31, 477–486.

Cohen, E.R., Feinglass, J., Barsuk, J.H., Barnard, C., O'Donnell, A., McGaghie, W.C., Wayne, D.B., 2010. Cost savings from reduced catheter-related bloodstream infection after

simulation-based education for residents in a medical intensive care unit. Simul. Healthc. J. Soc. Simul. Healthc. 5, 98–102. doi:10.1097/SIH.0b013e3181bc8304
Cohen, L., Manion, L., Morrison, K., Morrison, K.R.B., 2007. Research methods in education. Psychology Press.
Cook, D.A., 2010. One drop at a time: research to advance the science of simulation. Simul. Healthc. 5, 1–4.
Cook, D.A., Brydges, R., Hamstra, S.J., Zendejas, B., Szostek, J.H., Wang, A.T., Erwin, P.J., Hatala, R., 2012. Comparative effectiveness of technology-enhanced simulation versus other instructional methods: a systematic review and meta-analysis. Simul. Healthc. 7, 308–320.
Cook DA, H.R., 2011. Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. JAMA J. Am. Med. Assoc. 306, 978–988. doi:10.1001/jama.2011.1234
Cooper, J.G., Rutherford, E., Hamer, W., 2012. Survey of trainee attitudes to preparation for and participation in the FCEM examination. Emerg. Med. J. emermed–2011.
Coulter, A., 2002. After Bristol: putting patients at the centre. Qual. Saf. Health Care 11, 186–188. doi:10.1136/qhc.11.2.186
Cranton, P., 2001. Becoming an Authentic Teacher in Higher Education. Professional Practices in Adult Education and Human Resource Development Series. ERIC.
Crolla, R.M., van Ramshorst, B., Jansen, A., 1997. [Complication rate in laparoscopic cholecystectomy not different for residents in training and surgeons]. Ned. Tijdschr. Geneeskd. 141, 681–685.
Crossley, J., Humphris, G., Jolly, B., 2002. Assessing health professionals. Med. Educ. 36, 800–804.
Darling-Hammond, L., Snyder, J., 2000. Authentic assessment of teaching in context. Teach. Teach. Educ. 16, 523–545.
Dent, J.A., 2001. Current trends and future implications in the developing role of clinical skills centres. Med. Teach. 23, 483–489.
Department of Health, 2008. CMO Annual Report 2008 - Safer medical practice: Machines, manikins and polo mints [WWW Document]. URL http://www.dh.gov.uk/en/MediaCentre/Media/DH_096278# (accessed 3.26.10).
De Vries, A.M.M., Roten, Y., Meystre, C., Passchier, J., Despland, J.-N., Stiefel, F., 2014. Clinician characteristics, communication, and patient outcome in oncology: a systematic review. Psychooncology. 23, 375–381.
Dieckmann, P., Gaba, D., Rall, M., 2007. Deepening the theoretical foundations of patient simulation as social practice. Simul. Healthc. 2, 183.
Dieckmann, P., Manser, T., Wehner, T., Rall, M., 2007. Reality and Fiction Cues in Medical Patient Simulation: An Interview Study With Anesthesiologists. J. Cogn. Eng. Decis. Mak. 1, 148–168. doi:10.1518/155534307X232820
Downing, S.M., 2003. Validity: on the meaningful interpretation of assessment data. Med. Educ. 37, 830–837.
Dunkin, B., Adrales, G.L., Apelgren, K., Mellinger, J.D., 2007. Surgical simulation: a current review. Surg. Endosc. 21, 357–366.
Ellaway, R.H., Kneebone, R., Lachapelle, K., Topps, D., 2009. Practica continua: Connecting and combining simulation modalities for integrated teaching, learning and assessment. Med. Teach. 31, 725–731. doi:10.1080/01421590903124716

Epstein, R.M., Hundert, E.M., 2002. Defining and Assessing Professional Competence. JAMA 287, 226–235. doi:10.1001/jama.287.2.226

Ericsson, K.A., 2007. An expert-performance perspective of research on medical expertise: the study of clinical performance. [WWW Document]. Med. Educ. URL http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=18045365

Eversbusch, A., Grantcharov, T.P., 2004. Learning curves and impact of psychomotor training on performance in simulated colonoscopy: a randomized trial using a virtual reality endoscopy trainer. Surg. Endosc. Interv. Tech. 18, 1514–1518.

Fanning, R.M., Gaba, D.M., 2007. The role of debriefing in simulation-based learning. Simul. Healthc. J. Soc. Simul. Healthc. 2, 115–25. doi:10.1097/SIH.0b013e3180315539

Feilzer, M.Y., 2010. Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. J. Mix. Methods Res. 4, 6–16.

Fidment, S., 2012. The objective structured clinical exam (OSCE): a qualitative study exploring the healthcare student's experience. Stud. Engagem. Exp. J. 1.

Flin, R., Patey, R., Glavin, R., Maran, N., 2010. Anaesthetists' non-technical skills. Br. J. Anaesth. 105, 38–44.

Francis, R., 2013. The Mid Staffordshire NHS Foundation Trust Public Inquiry—Chaired by Robert Francis QC. Rep. Staffs. NHS Found. Trust Public Inq. Exec. Summ. Lond. TSO.

Frontier Economic, 2015. Exploring the costs of unsafe care in the NHS: a report prepared for the department of health, Frontier Economics, London. 2014.

Frymier, A.B., Shulman, G.M., 1995. "What's in it for me?": Increasing content relevance to enhance students' motivation. Commun. Educ. 44, 40–50.

Gaba, D.M., 2004. The future vision of simulation in health care. Qual. Saf. Health Care 13, i2–i10. doi:10.1136/qshc.2004.009878

Gaba, D.M., Howard, S.K., Fish, K.J., Smith, B.E., Sowb, Y.A., 2001. Simulation-based training in anesthesia crisis resource management (ACRM): a decade of experience. Simul. Gaming 32, 175.

Gaca, A.M., Frush, D.P., Hohenhaus, S.M., Luo, X., Ancarana, A., Pickles, A., Frush, K.S., 2007. Enhancing pediatric safety: using simulation to assess radiology resident preparedness for anaphylaxis from intravenous contrast media. Radiology 245, 236–244. doi:10.1148/radiol.2451061381

Gaskin, P.R., Owens, S.E., Talner, N.S., Sanders, S.P., Li, J.S., 2000. Clinical auscultation skills in pediatric residents. Pediatrics 105, 1184–1187.

Girzadas, D.V.J., Antonis, M.S., Zerth, H., Lambert, M., Clay, L., Bose, S., Harwood, R., 2009. Hybrid simulation combining a high fidelity scenario with a pelvic ultrasound task trainer enhances the training and evaluation of endovaginal ultrasound skills. [WWW Document]. Acad. Emerg. Med. Off. J. Soc. Acad. Emerg. Med. URL http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=19388924

Godden, D.R., Baddeley, A.D., 1975. Context-dependent memory in two natural environments: On land and underwater. Br. J. Psychol. 66, 325–331.

Gorter, S., Rethans, J., Van Der Heijde, D., Scherpbier, A., Houben, H., Van Der Vleuten, C., Van Der Linden, S., 2002. Reproducibility of clinical performance assessment in practice using incognito standardized patients. Med. Educ. 36, 827–832. doi:10.1046/j.1365-2923.2002.01296.x

Gould, D.A., Kessel, D.O., Healey, A.E., Johnson, S.J., Lewandowski, W.E., 2006. Simulators in catheter-based interventional radiology: training or computer games? Clin. Radiol. 61, 556–561. doi:10.1016/j.crad.2006.01.013

Greenaway, D., 2013. Securing the future of excellent patient care: Final report of the independent review. Shape of Training.

Gregg, M.J., Clark, T., 2007. Theoretical and practical applications of mental imagery, in: International Symposium on Performance Science. pp. 295–300.

Grunwald, T., Clark, D., Fisher, S., McLaughlin, M., Narayanan, S., Piepol, D., 2004. Using cognitive task analysis to facilitate collaboration in development of simulator to accelerate surgical training. Stud. Health Technol. Inform. 98, 114.

Gulikers, J., Bastiaens, T., Kirschner, P., 2004. A five-dimensional framework for authentic assessment. Educ. Technol. Res. Dev. 52, 67–86. doi:10.1007/BF02504676

Gulikers, J.T.M., Bastiaens, T.J., Martens, R.L., 2005. The surplus value of an authentic learning environment. Comput. Hum. Behav. 21, 509–521. doi:10.1016/j.chb.2004.10.028

Hammick, M., Dornan, T., Steinert, Y., 2010. Conducting a best evidence systematic review. Part 1: From idea to data coding. BEME Guide No. 13. Med. Teach. 32, 3–15. doi:10.3109/01421590903414245

Hancock, P.A., Vincenzi, D.A., Wise, J.A., Mouloua, M., 2008. Human factors in simulation and training. CRC Press.

Harden, R.M., 1988. What is an OSCE? Med. Teach. 10, 19–22.

Harden, R.M., Stevenson, M., Downie, W.W., Wilson, G.M., 1975. Assessment of clinical competence using objective structured examination. Br Med J 1, 447–451. doi:10.1136/bmj.1.5955.447

Hastorf, A.H., Cantril, H., 1954. They saw a game; a case study. J. Abnorm. Soc. Psychol. 49, 129–134. doi:10.1037/h0057880

Herrington, J., Reeves, T.C., Oliver, R., 2014. Authentic learning environments. Springer.

Higham, J., Nestel, D., Lupton, M., Kneebone, R., 2007. Teaching and learning gynaecology examination with hybrid simulation. Clin. Teach. 4, 238–243. doi:10.1111/j.1743-498X.2007.00179_1.x

Hinde, T., Gale, T., Anderson, I., Roberts, M., Sice, P., 2016. A study to assess the influence of interprofessional point of care simulation training on safety culture in the operating theatre environment of a university teaching hospital. J. Interprof. Care 1–3.

Hodges, B., 2006. Medical education and the maintenance of incompetence. Med. Teach. 28, 690–696. doi:10.1080/01421590601102964

Hodges, B., 2003. Validity and the OSCE. Med. Teach. 25, 250–254.

Hodges, B., Regehr, G., McNaughton, N., Tiberius, R., Hanson, M., 1999. OSCE checklists do not capture increasing levels of expertise. Acad. Med. 74, 1129.

Horwitz, L.I., 2011. Why have working hour restrictions apparently not improved patient safety? BMJ 342, d1200–d1200. doi:10.1136/bmj.d1200

Howe, A., Campion, P., Searle, J., Smith, H., 2004. New perspectives—approaches to medical education at four new UK medical schools. Bmj 329, 327–331.

Howe, K.R., 1988. Against the Quantitative-Qualitative Incompatibility Thesis or Dogmas Die Hard. Educ. Res. 17, 10 –16. doi:10.3102/0013189X017008010

Hull, D., 1993. Opening Minds, Opening Doors: The Rebirth of American Education. Center for Occupational Research and Development, PO Box 21206, Waco, TX 76702-1206.

Hunt, P.A., Lord, S.R., Taylor, S.C., 2004. Preparing for the MFAEM Examination. J. R. Army Med. Corps 150, 137–144.

Ilangaratne, J., 2011. Fundamental purpose of EWTD. BMJ 342, d2421–d2421. doi:10.1136/bmj.d2421

Isenberg, G.A., Berg, K.W., Veloski, J.A., Berg, D.D., Veloski, J.J., Yeo, C.J., 2011. Evaluation of the use of patient-focused simulation for student assessment in a surgery clerkship. Am. J. Surg. 201, 835–840. doi:10.1016/j.amjsurg.2010.01.034

Issenberg, S.B., Mcgaghie, W.C., Petrusa, E.R., Lee Gordon, D., Scalese, R.J., 2005. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med. Teach. 27, 10–28. doi:10.1080/01421590500046924

Issenberg, S.B., Ringsted, C., Østergaard, D., Dieckmann, P., 2011. Setting a Research Agenda for Simulation-Based Healthcare Education. Simul. Healthc. J. Soc. Simul. Healthc. 6, 155–167. doi:10.1097/SIH.0b013e3182207c24

Jarvis, P., Parker, S., 2006. Human Learning: An Holistic Approach. Routledge.

Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed Methods Research: A Research Paradigm Whose Time Has Come. Educ. Res. 33, 14–26. doi:10.3102/0013189X033007014

Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A., 2007. Toward a definition of mixed methods research. J. Mix. Methods Res. 1, 112–133.

Jonassen, D., Davidson, M., Collins, M., Campbell, J., Haag, B.B., 1995. Constructivism and computer-mediated communication in distance education. Am. J. Distance Educ. 9, 7–26.

Karwowski, W., 2006. International Encyclopedia of Ergonomics and Human Factors, Second Edition - 3 Volume Set. CRC Press.

Kassab, E., Tun, J.K., Arora, S., King, D., Ahmed, K., Miskovic, D., Cope, A., Vadhwana, B., Bello, F., Sevdalis, N., Kneebone, R., 2011. "Blowing up the Barriers" in Surgical Training. Ann. Surg. 1. doi:10.1097/SLA.0b013e318228944a

Kaufman, D.M., Mann, K.V., 2000. Teaching and learning in medical education: how theory can inform practice. Underst. Med. Educ. 16–36.

Kenney, P.A., Wszolek, M.F., Gould, J.J., Libertino, J.A., Moinzadeh, A., 2009. Face, content, and construct validity of dV-trainer, a novel virtual reality simulator for robotic surgery. Urology 73, 1288–1292.

Ker, J., Hogg, G., Maran, N., Walsh, K., 2010. Cost-effective simulation. Cost Eff. Med. Educ. 61–71.

Kim, J.Y., Myung, S.J., 2014. Could clinical experience during clerkship enhance students' clinical performance? BMC Med. Educ. 14, 209.

Kitzinger, J., 1995. Qualitative research: introducing focus groups. Bmj 311, 299–302.

Kneebone, R., 2005. Evaluating clinical simulations for learning procedural skills: a theory-based approach. Acad. Med. 80, 549–553.

Kneebone, R., 2002. Total internal reflection: an essay on paradigms. Med. Educ. 36, 514–518.

Kneebone, R.L., 2009. Practice, rehearsal, and performance: an approach for simulation-based surgical and procedure training. JAMA 302, 1336–1338.

Kneebone, R.L., D Nestel, C Vincent, A Darzi, 2007. Complexity, risk and simulation in learning procedural skills. Med. Educ. 41, 808–814. doi:10.1111/j.1365-2923.2007.02799.x

Kneebone, R.L., Kidd, J., Nestel, D., Barnet, A., Lo, B., King, R., Yang, G.Z., Brown, R., 2005. Blurring the boundaries: scenario-based simulation in a clinical setting. Med. Educ. 39, 580–587.

Kneebone, R.L., W Scott, A Darzi, M Horrocks, 2004. Simulation and clinical practice: strengthening the relationship. Med. Educ. 38, 1095–1102. doi:10.1111/j.1365-2929.2004.01959.x

Kneebone, R., Nestel, D., Wetzel, C., Black, S., Jacklin, R., Aggarwal, R., Yadollahi, F., Wolfe, J., Vincent, C., Darzi, A., 2006. The human face of simulation: patient-focused simulation training. Acad. Med. J. Assoc. Am. Med. Coll. 81, 919–924. doi:10.1097/01.ACM.0000238323.73623.c2

Kneebone, R., Nestel, D., Yadollahi, F., Brown, R., Nolan, C., Durack, J., Brenton, H., Moulton, C., Archer, J., Darzi, A., 2006. Assessing procedural skills in context: exploring the feasibility of an Integrated Procedural Performance Instrument (IPPI). Med. Educ. 40, 1105–1114. doi:10.1111/j.1365-2929.2006.02612.x

Kohn, L.T., Corrigan, J., Donaldson, M.S., America, I. of M. (U. S.). C. on Q. of H.C. in, 2000. To err is human: building a safer health system. National Academies Press.

Kolb, A.Y., Kolb, D.A., 2012. Experiential learning theory, in: Encyclopedia of the Sciences of Learning. Springer, pp. 1215–1219.

Kolb, D.A., 1984. Experiential Learning: Experience as the Source of Learning and Development. Financial Times/ Prentice Hall.

Krauss, S.E., 2005. Research paradigms and meaning making: A primer. Qual. Rep. 10, 758–770.

Kreber, C., Klampfleitner, M., 2012. Construing the meaning of authenticity in university teaching: Comparing explicit to implicit theories. J. Constr. Psychol. 25, 34–69.

Kwasnicki, R.M., Lewis, T.M., Reissis, D., Sarvesvaran, M., Paraskeva, P.A., 2012. A high fidelity model for single-incision laparoscopic cholecystectomy. Int. J. Surg. Lond. Engl. 10, 285–289. doi:10.1016/j.ijsu.2012.04.015

Kyaw Tun, J., Alinier, G., Tang, J., Kneebone, R., In Press. Redefining simulation fidelity for healthcare education. Simul. Gaming.

Kyaw Tun, J., Alinier, G., Tang, J., Kneebone, R.L., 2015. Redefining Simulation Fidelity for Healthcare Education. Simul. Gaming 1046878115576103.

Kyaw Tun, J., Granados, A., Mavroveli, S., Nuttall, S., Kadiyala, A.N., Brown, R., Bello, F., Kneebone, R.L., 2012. Simulating Various Levels of Clinical Challenge in the Assessment of Clinical Procedure Competence. Ann. Emerg. Med. 60, 112–120.e5. doi:10.1016/j.annemergmed.2012.01.036

Kyaw Tun, J., Kneebone, R., 2011. Bridging worlds: applying the science of motor learning to clinical education. Med. Educ. 45, 111–114. doi:10.1111/j.1365-2923.2010.03914.x

Kyaw Tun, J., Tang, J., Kneebone, R., 2011. Redefining simulation fidelity for healthcare education. Presented at the Society in Europe for Simulation Applied to Medicine, Granada, Spain.

Lane, C., Rollnick, S., 2007. The use of simulated patients and role-play in communication skills training: a review of the literature to August 2005. Patient Educ. Couns. 67, 13–20.

Lane, N.E., Alluisi, E.A., 1992. Fidelity and Validity in Distributed Interactive Simulation: Questions and Answers.

Laschinger, S., Medves, J., Pulling, C., McGraw, D.R., Waytuck, B., Harrison, M.B., Gambeta, K., 2008. Effectiveness of simulation on health profession students' knowledge, skills, confidence and satisfaction. Int. J. Evid. Based Healthc. 6, 278–302. doi:10.1111/j.1744-1609.2008.00108.x

Lave, J., Wenger, E., 1991. Situated Learning: Legitimate Peripheral Participation. Cambridge University Press.

Laverty, S.M., 2003. Hermeneutic phenomenology and phenomenology: A comparison of historical and methodological considerations. Int. J. Qual. Methods 2, 21–35.

Leape, L.L., Woods, D.D., Hatlie, M.J., Kizer, K.W., Schroeder, S.A., Lundberg, G.D., 1998. Promoting Patient Safety by Preventing Medical Error. JAMA 280, 1444–1447. doi:10.1001/jama.280.16.1444

LeBlanc, V.R., Tabak, D., Kneebone, R., Nestel, D., MacRae, H., Moulton, C.-A., 2009. Psychometric properties of an integrated assessment of technical and communication skills. Am. J. Surg. 197, 96–101.

Leech, N., Onwuegbuzie, A., 2009. A typology of mixed methods research designs. Qual. Quant. 43, 265–275. doi:10.1007/s11135-007-9105-3

Leuba, C., Lucas, C., 1945. The effects of attitudes on descriptions of pictures. J. Exp. Psychol. 35, 517–524. doi:10.1037/h0056924

Leung, W.-C., Diwakar, V., 2002. Competency based medical training: review * Commentary: The baby is thrown out with the bathwater. BMJ 325, 693–696. doi:10.1136/bmj.325.7366.693

Levett-Jones, T., Lathlean, J., 2008. Belongingness: A prerequisite for nursing students' clinical learning. Nurse Educ. Pract. 8, 103–111.

Lindquist, E., 1951. Preliminary considerations in objective test construction. Educ. Meas. 119–158.

Lingard, L., Reznick, R., DeVito, I., Espin, S., 2002. Forming professional identities on the health care team: discursive constructions of the "other" in the operating room. Med. Educ. 36, 728–734. doi:10.1046/j.1365-2923.2002.01271.x

Lombardi, M.M., 2007. Authentic learning for the 21st century: An overview.

Lynagh, M., Burton, R., Sanson-Fisher, R., 2007. A systematic review of medical skills laboratory training: where to from here? Med. Educ. 41, 879–887. doi:10.1111/j.1365-2923.2007.02821.x

Mackenzie, D., 1973. The history of sutures. Med. Hist. 17, 158–168.

Mangione, S., Nieman, L.Z., 1997. Cardiac Auscultatory Skills of Internal Medicine and Family Practice Trainees. JAMA J. Am. Med. Assoc. 278, 717 –722. doi:10.1001/jama.1997.03550090041030

Maran, N.J., Glavin, R.J., 2003. Low- to high-fidelity simulation - a continuum of medical education? Med. Educ. 37 Suppl 1, 22–28.

Marks, D.F., 2010. IQ variations across time, race, and nationality: an artifact of differences in literacy skills. Psychol. Rep. 106, 643–664. doi:10.2466/PR0.106.3.643-664

Marsden, A., 2014. Empathetic Consultation Skills in Undergraduate Medical Education: A Qualitative Approach. University of East Anglia.

Martin, I.G., Stark, P., Jolly, B., 2000. Benefiting from clinical experience: the influence of learning style and clinical experience on performance in an undergraduate objective structured clinical examination. Med. Educ. 34, 530–534.

Martin, J.A., Regehr, G., Reznick, R., Macrae, H., Murnaghan, J., Hutchison, C., Brown, M., 1997. Objective structured assessment of technical skill (OSATS) for surgical residents. Br. J. Surg. 84, 273–278. doi:10.1002/bjs.1800840237

Martin, P., Bateson, P.P.G., 1993. Measuring behaviour: an introductory guide. Cambridge University Press.

Maxwell, J.A., 2008. Designing a qualitative study. Sage Handb. Appl. Soc. Res. Methods 214–253.

McDougall, E.M., Corica, F.A., Boker, J.R., Sala, L.G., Stoliar, G., Borin, J.F., Chu, F.T., Clayman, R.V., 2006. Construct validity testing of a laparoscopic surgical simulator. J. Am. Coll. Surg. 202, 779–787.

McGaghie, W.C., Issenberg, S.B., Cohen, E.R., Barsuk, J.H., Wayne, D.B., 2011. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. Acad. Med. 86, 706.

Mcgaghie, W.C., Issenberg, S.B., Petrusa, E.R., Scalese, R.J., 2010. A critical review of simulation-based medical education research: 2003-2009. Med. Educ. 44, 50–63. doi:10.1111/j.1365-2923.2009.03547.x

McIlrath, D., Huitt, W., 1995. The teaching-learning process: A discussion of models. Educ. Psychol. Interact.

McKenzie Smith, M., Turkhud, K., 2013. Simulation-based education in support of HCA development. Br. J. Healthc. Assist. 7.

McKinney, J., Cook, D.A., Wood, D., Hatala, R., 2013. Simulation-based training for cardiac auscultation skills: systematic review and meta-analysis. J. Gen. Intern. Med. 28, 283–291. doi:10.1007/s11606-012-2198-y

Mehdi, Z., Roots, A., Ernst, T., Birns, J., Ross, A., Reedy, G., Jaye, P., 2014. Simulation training for geriatric medicine. Clin. Teach. 11, 387–392.

Mehrens, W.A., 1992. Using Performance Assessment for Accountability Purposes. Educ. Meas. Issues Pract. 11, 3–9. doi:10.1111/j.1745-3992.1992.tb00220.x

Messick, S., 1994. The interplay of evidence and consequences in the validation of performance assessments. Educ. Res. 23, 13.

M. Harden, R., 1999. BEME Guide No. 1: Best evidence medical education. Med. Teach. 21, 553–562.

Millenson, M.L., 2002. Pushing the profession: how the news media turned patient safety into a priority. Qual. Saf. Health Care 11, 57 –63. doi:10.1136/qhc.11.1.57

Montgomery, K., 2002. Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. Coll. Teach. 50, 34–40.

Morgan, G., Smircich, L., 1980. The case for qualitative research. Acad. Manage. Rev. 491–500.

Moulton, C., Tabak, D., Kneebone, R., Nestel, D., MacRae, H., LeBlanc, V.R., 2009. Teaching communication skills using the integrated procedural performance instrument (IPPI): A randomized controlled trial. Am. J. Surg. 197, 113–118.

Mueller, J., 2008. What is Authentic Assessment?'Authentic Assessment Toolbox.

Nestel, D., Kneebone, R., Nolan, C., Akhtar, K., Darzi, A., 2009. Formative assessment of procedural skills: students' responses to the Objective Structured Clinical Examination and the Integrated Performance Procedural Instrument. Assess. Eval. High. Educ. doi:10.1080/02602930903221469

Nestel, D., Walker, K., Simon, R., Aggarwal, R., Andreatta, P., 2011. Nontechnical skills: an inaccurate and unhelpful descriptor? Simul. Healthc. 6, 2–3.

Newmann, F.M., Carmichael, D.L., King, M.B., 2015. Authentic Intellectual Work: Improving Teaching for Rigorous Learning. Corwin Press.

Newmann, F.M., Wehlage, G.G., 1993. Five standards of authentic instruction. Educ. Leadersh. 50, 8–12.
NHS Confederation, 2011. The impact of non-payment for acute re-admissions. Briefing 211, 1–10.
Nollaig, F., 2011. Qualitative Research Methods In Psychology: Combining Core Approaches: From core to combined approaches. McGraw-Hill Education (UK).
Norcini, J., Burch, V., 2007. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med. Teach. 29, 855–71. doi:10.1080/01421590701775453
O'Brien, B., Cooke, M., Irby, D.M., 2007. Perceptions and attributions of third-year student struggles in clerkships: do students and clerkship directors agree? Acad. Med. 82, 970–978.
Oddone, E.Z., Waugh, R.A., Samsa, G., Corey, R., Feussner, J.R., 1993. Teaching cardiovascular examination skills: results from a randomized controlled trial. Am. J. Med. 95, 389–396.
Okuda, Y., Bond, W., Bonfante, G., McLaughlin, S., Spillane, L., Wang, E., Vozenilek, J., Gordon, J.A., Gillett, B., Peckler, B., Sinert, R., Onkst, C., Nabors, S., Issley, S., Maguire, C., Galwankarm, S., Arquilla, B., 2008. Simulation in a disaster drill: comparison of high-fidelity simulators versus trained actors. [WWW Document]. Acad. Emerg. Med. Off. J. Soc. Acad. Emerg. Med. URL http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=18717651
Olympio, M.A., Whelan, R., Ford, R.P.A., Saunders, I.C.M., 2003. Failure of simulation training to change residents' management of oesophageal intubation. Br. J. Anaesth. 91, 312–318.
Pashler, H., McDaniel, M., Rohrer, D., Bjork, R., 2008. Learning styles concepts and evidence. Psychol. Sci. Public Interest 9, 105–119.
Perlini, S., Salinaro, F., Santalucia, P., Musca, F., 2014. Simulation-guided cardiac auscultation improves medical students' clinical skills: the Pavia pilot experience. Intern. Emerg. Med. 9, 165–172.
Ponton-Carss, A., Hutchison, C., Violato, C., 2011. Assessment of communication, professionalism, and surgical skills in an objective structured performance-related examination (OSPRE): a psychometric study. Am. J. Surg. 202, 433–440.
Quantrill, S.J., Kyaw Tun, J., 2012. Workplace-based assessment as an educational tool. Guide supplement 31.5 - Viewpoint. Med. Teach. 34, 417–418. doi:10.3109/0142159X.2012.668234
Quilici, A.P., Pogetti, R.S., Fontes, B., Zantut, L.F.C., Chaves, E.T., Birolini, D., 2005. Is the Advanced Trauma Life Support simulation exam more stressful for the surgeon than emergency department trauma care? Clin. São Paulo Braz. 60, 287–292. doi:/S1807-59322005000400005
Rashid, M.S., Sobowale, O., Gore, D., 2011. A near-peer teaching program designed, developed and delivered exclusively by recent medical graduates for final year medical students sitting the final objective structured clinical examination (OSCE). BMC Med. Educ. 11, 11.
RCR, 2012. The Royal College of Radiologists | Specialty Training Curriculum for Clinical Radiology May 2012 [WWW Document]. URL http://www.rcr.ac.uk/content.aspx?PageID=2067 (accessed 8.15.12).

Regan, K., Harney, L., Goodhand, K., Strath, A., Vosper, H., 2014. Pharmacy Simulation: A Scottish, Student-Led Perspective with Lessons for the UK and Beyond. Pharmacy 2, 50–64.

Remmen, R., Derese, A., Scherpbier, A., Denekens, J., Hermann, I., van der Vleuten, C., Van Royen, P., Bossaert, L., 1999. Can medical schools rely on clerkships to train students in basic clinical skills? Med. Educ. 33, 600–605.

Rethans, J.-J., Norcini, J.J., Baron-Maldonado, M., Blackmore, D., Jolly, B.C., LaDuca, T., Lew, S., Page, G.G., Southgate, L.H., 2002. The relationship between competence and performance: implications for assessing practice performance. Med. Educ. 36, 901–9.

Reznek, M., Smith-Coggins, R., Howard, S., Kiran, K., Harter, P., Sowb, Y., Gaba, D., Krummel, T., 2003. Emergency Medicine Crisis Resource Management (EMCRM): Pilot study of a simulation-based crisis management course for emergency medicine. Acad. Emerg. Med. 10, 386–389.

Riley, S., Kumar, N., 2012. Teaching medical professionalism. Clin. Med. 12, 9–11.

Roger Kneebone, Nestel, D., Wetzel, C., Black, S., Jacklin, R., Aggarwal, R., Yadollahi, F., Wolfe, J., Vincent, C., Darzi, A., 2006. The Human Face of Simulation: Patient-Focused Simulation Training. Acad. Med. 81, 919–924. doi:10.1097/01.ACM.0000238323.73623.c2

Roldan, C.A., Shively, B.K., Crawford, M.H., 1996. Value of the cardiovascular physical examination for detecting valvular heart disease in asymptomatic subjects. Am. J. Cardiol. 77, 1327–1331.

Rollason, C., Kyaw Tun, J., Gill, R., Smith, T., Kandasamy, G., 2009. Improving Fire Management and Policy in Anaesthetic Environments Through Collaborative Simulation Training. Presented at the Nation Association of Medical Simulators, Manchester, UK.

Royal College of Surgeons of England, 2014. Good Surgical Practice.

Rudland, J., Wilkinson, T., Smith-Han, K., Thompson-Fawcett, M., 2008. "You can do it late at night or in the morning. You can do it at home, I did it with my flatmate." The educational impact of an OSCE - PB - Informa Healthcare. Med. Teach. 30, 206. doi:10.1080/01421590701851312

Sachdeva, A.K., Pellegrini, C.A., Johnson, K.A., 2008. Support for simulation-based surgical education through American College of Surgeons--accredited education institutes. World J. Surg. 32, 196–207. doi:10.1007/s00268-007-9306-x

Sale, J.E.M., Lohfeld, L.H., Brazil, K., 2002. Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. Qual. Quant. 36, 43–53.

Sarwani, N., Tappouni, R., Flemming, D., 2012. Use of a Simulation Laboratory to Train Radiology Residents in the Management of Acute Radiologic Emergencies. Am. J. Roentgenol. 199, 244–251.

Scalese, R., Obeso, V., Issenberg, S., 2008. Simulation Technology for Skills Training and Competency Assessment in Medical Education. J. Gen. Intern. Med. 23, 46–49. doi:10.1007/s11606-007-0283-4

Schifferdecker, K.E., Reed, V.A., 2009. Using mixed methods research in medical education: basic guidelines for researchers. Med. Educ. 43, 637–644. doi:10.1111/j.1365-2923.2009.03386.x

Schlachta, C.M., Mamazza, J., Seshadri, P.A., Cadeddu, M., Gregoire, R., Poulin, E.C., 2001. Defining a learning curve for laparoscopic colorectal resections. Dis. Colon Rectum 44, 217–222. doi:10.1007/BF02234296

Schmidt, R.A., Lee, T., 1988. Motor control and learning. Human kinetics.

Schön, D.A., 1987. Educating the reflective practitioner: toward a new design for teaching and learning in the professions. Jossey-Bass.

Schout, B.M.A., Hendrikx, A.J.M., Scheele, F., Bemelmans, B.L.H., Scherpbier, A., 2010. Validation and implementation of surgical simulators: a critical review of present, past, and future. Surg. Endosc. 24, 536–546.

Schrag, F., 1992. In defense of positivist research paradigms. Educ. Res. 21, 5.

Schuwirth, L.W.T., van der Vleuten, C.P.M., 2003. The use of clinical simulations in assessment. Med. Educ. 37, 65–71. doi:10.1046/j.1365-2923.37.s1.8.x

Schuwirth, L.W., van der Vleuten, C.P., 2011. General overview of the theories used in assessment: AMEE Guide No. 57. Med. Teach. 33, 783–797.

Schwandt, T.A., 1994. Constructivist, interpretivist approaches to human inquiry.

Shah, P.R., Joseph, A., Haray, P.N., 2005. Laparoscopic colorectal surgery: learning curve and training implications. Postgrad. Med. J. 81, 537–540. doi:10.1136/pgmj.2004.028100

Shanthikumar, J.G., Sargent, R.G., 1983. A unifying view of hybrid simulation/analytic models and modeling. Oper. Res. 31, 1030–1052.

Siassakos, D., Draycott, T., O'Brien, K., Kenyon, C., Bartlett, C., Fox, R., 2010. Exploratory randomized controlled trial of hybrid obstetric simulation training for undergraduate students. Simul. Healthc. J. Soc. Simul. Healthc. 5, 193–198. doi:10.1097/SIH.0b013e3181d3ee0b

Sica, G., Barron, D., Blum, R., Frenna, T., Raemer, D., 1999. Computerized realistic simulation: a teaching module for crisis management in radiology. Am. J. Roentgenol. 172, 301.

Simon, E.L., Lecat, P.J., Haller, N.A., Williams, C.J., Martin, S.W., Carney, J.A., Pakiela, J.A., 2012. Improved Auscultation Skills in Paramedic Students Using a Modified Stethoscope. J. Emerg. Med. 43, 1091–1097. doi:10.1016/j.jemermed.2012.01.048

Smith, J.A., Flowers, P., Larkin, M., 2009. Interpretative Phenomenological Analysis: Theory, Method and Research. SAGE.

Spector, M.J., Merrill, M.D., Merrienboer, J. van, Driscoll, M.P., 2007. Handbook of Research on Educational Communications and Technology: A Project of the Association for Educational Communications and Technology. Routledge.

Splitter, L.J., 2008. Authenticity and Constructivism in Education. Stud. Philos. Educ. 28, 135–151. doi:10.1007/s11217-008-9105-3

Spruit, E.N., Band, G.P., Hamming, J.F., Ridderinkhof, K.R., 2014. Optimal training design for procedural motor skills: a review and application to laparoscopic surgery. Psychol. Res. 78, 878–891.

Starkes, J.L., Ericsson, K.A., 2003. Expert Performance in Sports: Advances in Research on Sport Expertise. Human Kinetics.

Sutherland, L.M., Middleton, P.F., Anthony, A., Hamdorf, J., Cregan, P., Scott, D., Maddern, G.J., 2006. Surgical Simulation: A Systematic Review. Ann. Surg. 243, 291–300. doi:10.1097/01.sla.0000200839.93965.26

Tavakol, M., 2009. Medical Research Paradigms. J. Med. Educ. 5.

Terwilliger, J., 1997. Research news and Comment: Semantics, Psychometrics, and Assessment Reform: A Close Look at “Authentic” Assessments. Educ. Res. 26, 24–27. doi:10.3102/0013189X026008024

The Royal College of Emergency Medicine, n.d. Simulaton Training [WWW Document]. URL http://www.rcem.ac.uk/Training-Exams/Training/Simulation%20Training (accessed 3.14.16).

The Royal College Of Radiologists, 2010. Simulation in radiology.

Thompson, W.R., Hayek, C.S., Tuchinda, C., Telford, J.K., Lombardo, J.S., 2001. Automated cardiac auscultation for detection of pathologic heart murmurs. Pediatr. Cardiol. 22, 373–379.

Tubbs, R.J., Murphy, B., Mainiero, M.B., Shapiro, M., Kobayashi, L., Lindquist, D., Smith, J.L., Siegel, N., 2009. High-fidelity medical simulation as an assessment tool for radiology residents’ acute contrast reaction management skills. J. Am. Coll. Radiol. JACR 6, 582–587. doi:10.1016/j.jacr.2009.04.012

van der Vleuten, C.P.M., Schuwirth, L.W.T., 2005. Assessing professional competence: from methods to programmes. Med. Educ. 39, 309–17. doi:10.1111/j.1365-2929.2005.02094.x

van de Vleuten, C.P.M., 1996. The assessment of professional competence: Developments, research and practical implications. Adv. Health Sci. Educ. 1, 41–67. doi:10.1007/BF00596229

Våpenstad, C., Hofstad, E.F., Langø, T., Mårvik, R., Chmarra, M.K., 2013. Perceiving haptic feedback in virtual reality simulators. Surg. Endosc. 27, 2391–2397. doi:10.1007/s00464-012-2745-y

Verma, A., Bhatt, H., Booton, P., Kneebone, R., 2011. The Ventriloscope® as an innovative tool for assessing clinical examination skills: Appraisal of a novel method of simulating auscultatory findings. Med. Teach. 33, 388–396.

Wass, V., Van Der Vleuten, C., 2004. The long case. Med. Educ. 38, 1176–1180.

Wass, V., Van der Vleuten, C., Shatzer, J., Jones, R., 2001. Assessment of clinical competence. The Lancet 357, 945–949. doi:10.1016/S0140-6736(00)04221-5

Weinger, M.B., 2010. The pharmacology of simulation: a conceptual framework to inform progress in simulation research. Simul. Healthc. 5, 8–15.

Weller, J.M., 2004. Simulation in undergraduate medical education: bridging the gap between theory and practice. [WWW Document]. Med. Educ. URL http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=14962024

Wenger, E., 2011. Communities of practice: A brief introduction.

Wenger, E., 2000. Communities of practice and social learning systems. Organization 7, 225–246.

Wenk, M., Waurick, R., Schotes, D., Wenk, M., Gerdes, C., Aken, H.K.V., Pöpping, D.M., 2009. Simulation-based medical education is no better than problem-based discussions and induces misjudgment in self-assessment. Adv. Health Sci. Educ. 14, 159–171. doi:10.1007/s10459-008-9098-2

Wiggins, G., 1993. Assessment: Authenticity, Context, and Validity. Phi Delta Kappan 75, 200–214.

Wiggins, G., 1990. The Case for Authentic Assessment. ERIC Digest.

Wiggins, G., 1989. Teaching to the (Authentic) Test. Educ. Leadersh. 46, 41–47.

Wood, D., Bruner, J.S., Ross, G., 1976. The role of tutoring in problem solving*. J. Child Psychol. Psychiatry 17, 89–100.

Woodrum, D.T., Andreatta, P.B., Yellamanchilli, R.K., Feryus, L., Gauger, P.G., Minter, R.M., 2006. Construct validity of the LapSim laparoscopic surgical simulator. Am. J. Surg. 191, 28–32.

Wormald, B.W., Schoeman, S., Somasunderam, A., Penn, M., 2009. Assessment drives learning: An unavoidable truth? Anat. Sci. Educ. 2, 199–204. doi:10.1002/ase.102

Yardley, L., 2000. Dilemmas in qualitative health research. Psychol. Health 15, 215–228.

Yule, S., Flin, R., Paterson-Brown, S., Maran, N., 2006. Non-technical skills for surgeons in the operating room: a review of the literature. Surgery 139, 140–9. doi:10.1016/j.surg.2005.06.017

Zendejas, B., Wang, A.T., Brydges, R., Hamstra, S.J., Cook, D.A., 2013. Cost: the missing outcome in simulation-based medical education research: a systematic review. Surgery 153, 160–176.

Ziv, A., Wolpe, P.R., Small, S.D., Glick, S., 2003. Simulation-based medical education: an ethical imperative. Acad. Med. J. Assoc. Am. Med. Coll. 78, 783–788.

]

Appendices:

1) Participant Consent form for Study in Chapter 4 and 5

2) Participant Briefings for Simulation Scenarios

3) Simulated Patient Roles for Study in Chapter 4 and 5

4) Performance Rating Tools: OSATS TSC, OSATS GRS, DOPS Tool

5) Post-Simulation Questionnaire for PTT-Based and PFHS Scenarios

6) Sample Anonymised Interview Transcript

7) Participant Consent form for Study in Chapter 6

8) Tutorial Slides for Cardiovascular Examination Skills Course

9) OSCE Checklist and GRS

10) Sample Excerpt Anonymised Focus Groups Transcript

11) Sample Excerpt Responses to Pre and Post-Course Questionnaire on Knowledge and Understanding of Cardiovascular Examination.

12) Excerpt from Account of Observations

Appendix 1: Participant Information and Consent Form

Imperial College London

Participant Information Sheet
Hybrid Simulation Study

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?
Simulation is playing an increasing role in assessment of doctors' competence. We are investigating the validity of using hybrid simulation to assessment competence in clinical procedural skills.

Do I have to take part?
It is up to you whether or not to take part. If you decide to take part, you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason.

What will happen to me if I take part?
1. You will be asked to participate in a bench-top wound simulation scenario and two hybrid simulation scenario based on managing a wound
2. You will be asked to fill out questionnaires and answer questions in an interview.
3. Your performance in the scenarios will be recorded and rated for data analysis.

What are the possible disadvantages and risks of taking part?
Parts of the simulations will involve handling safe instruments and therefore we request that you practice with the accepted high standards of safety as expected in the real workplace. Otherwise we do not anticipate any disadvantages.

What are the possible benefits to taking part?
The possible benefits of this study are to create valid authentic and more purpose-fit assessments for procedural skill competence as well as increasing our understanding of the use of simulation for purpose of assessments.

What if something goes wrong?
We do not foresee anything going wrong. However if there are any problems, you are free to terminate the simulation or post-simulation questionnaire/interviews at any time without giving any reason.

Will my taking part in this study be kept confidential?
All the data we collect from you will be anonymised by allocating you a code. Any paperwork containing your identity (consent form) will be kept separate to anonymised data collected during the study. They will not be shown to anyone outside the research team and can be destroyed as soon as your participation in the project comes to an end if you request.

What will happen to the results of the research study?
The results will be presented in departmental meetings, international conferences and peer reviewed journals. All data presented will be anonymous and never related to any specific individual or team.

Who is funding the research?
This research is funded by the London Deanery.

Who has reviewed the study?
This study was given a favourable ethical opinion for conduct by St Mary's Research Ethics Committee.

Contact for further information
For more information, please contact:
Jimmy Kyaw Tun email: (Lead researcher)
Roger Kneebone email: (Research Supervisor)

Imperial College
London

Participant Identification Number:_________________

Consent Form

Study Title: Use of Patient-Focused Hybrid Simulation for Procedural Skills Learning and Assessment

Principle Researchers: Dr. Jimmy Kyaw Tun, Eva Kassab, Dr. Roger Kneebone, Dr Ruth Brown

PLEASE READ AND COMPLETE THE CONSENT FORM THOROUGHLY

- I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions ☐
- I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without legal rights being affected ☐
- I understand that my anonymised evaluations will be looked at by responsible individuals from our research team at Imperial College and hat my data will be collated with other participants' data so that I cannot be identified individually ☐
- I understand that my performance in the simulations will be rated by responsible individuals ☐
- I understand the principles of sharps instruments handling and have been cleared by occupational health by my NHS trust to take part is exposure prone procedures ☐
- I agree to take part in the above study ☐

________________________	_______________________	____/____/_____
Name of participant	Signature	Date

Appendix 2: Briefing Instructions to Participants

A) PFHS Simulation Scenario 1 (Simple)

You are an emergency medicine doctor asked to see a patient in the ED by the ED nurse to assess and manage a wound.
Female patient sustained a laceration to her **Left** thigh whilst cycling to work. The ED nurse thinks the wound needs closing and has asked you if you could assess and proceed. As the surgeon/emergency medicine doctor you will be required to:

- **Manage** the patient as you would in your workplace
- **Assess** the wound and close wound if appropriate

Equipment necessary for managing the patient is available.
A ED nurse is available to assist you.

B) PFHS Simulation Scenario 2 (Complex)

You are an emergency medicine doctor asked to see a patient in the ED by the ED nurse to assess and manage a wound.
Male patient sustained a laceration to his **Left** upper arm whilst travelling back home from a night out in town. The ED nurse thinks the wound needs closing and has asked you if you could assess and proceed.
As the surgeon/emergency practitioner you will be required to:

- **Manage** the patient as you would in your workplace
- **Assess** the wound and close wound

Equipment necessary for managing the patient is available.

An ED nurse is available to assist you.

Appendix 3 – Patient Roles and Instructions for Patient Focused Hybrid Simulations

A) Patient Focused Hybrid Simulation Scenario 1 - Instruction to Simulated Patient

Introduction

Your role has been designed for the actors playing the role of a simulated patient with a wound as part of a simulated patient encounter for wound management. This forms part of a study looking at the use of these scenarios for assessing doctor's competence.

Scenario Objectives

The aim of the scenario is to provide a realistic patient encounter of a traumatic laceration in which a trainee doctor can demonstrate and be assessed in a range of skills required to managed and close a wound. These include:

- Communication with patient and staff
- Technical skills
- Clinical and Equipment Knowledge
- Professionalism

Expected time to complete scenario is 20 minutes.

Background/ Profile

Name: Amy Cook
Age: 30-40s
Works in the city as a
Personal Background

- Lives with boyfriend. No children.
- Fit and well, no allergies or medical issues
- Had tetanus jab in the past when u cut finger (whilst chopping food) 5-6 years ago

Settings

The scenario takes place in a simulated A&E cubicle. All equipment necessary to complete the scenario is available.

Persons present include:

1. Participant Trainee Doctor (being assessed)
2. Simulated A&E nurse
3. Simulated patient with traumatic wound/laceration

Scenario Sequence

Prior to the scenario, you were cycling to work, and fell of your bicycle as you attempted to avoid a passing vehicle when you sustained a cut to your leg. Aside from the laceration, you have no other injuries.
You are sat on a bed with some gauze lightly covering your wound which the triage nurse applied.
The nurse has performed a basic set of "obs" – the patient's blood pressure (115/70mmhg) pulse (68bpm) and oxygen saturations (99% on air), which are normal.
The participant doctor should:

- Examine the wound
- Ask some relevant history and tetanus
- Irrigate the wound
- Draw up and Apply local anaesthetic
- Open suture pack
- Suture wound

Throughout the scenario, the participant should be interacting in a professional manner with the nurse and patient.
On completion of suturing, the nurse will tell the doctor that she will sort out the dressings and that she can get on with the documentations. This will be the signal to end the scenario.
During the scenario you should:

- Ask about the "injection" and suture to see if it is necessary

- Ask about scarring
- Ask about taking out sutures

Dress, Behaviour, Affect and Mannerisms of the Simulated Patient

- You are dressed in work clothes
    - Wear a skirt/shorts for easy access to wound
- You are friendly and co-operative but a little nervous.
- You flinch when you see the needles and when the needle punctures your skin (until you are anaesthetised/ numbed)
- If it does not appear that the doctor put enough local anaesthetic then show pain when he/she attempts to close the wound.
- You are happy to talk things like family, social life, how the workday went etc.

B) Patient Focused Hybrid Simulation Scenario 2 - Instruction to Simulated Patient

Introduction

Your role has been designed for the actors playing the role of a simulated patient with a wound as part of a simulated patient encounter for wound management. This forms part of a study looking at the use of these scenarios for assessing doctor's competence.

Scenario Objectives

The aim of the scenario is to provide a realistic patient encounter of a traumatic laceration in which a trainee doctor can demonstrate and be assessed in a range of skills required to managed and close a wound. These include:

- Communication with patient and staff
- Technical skills
- Clinical and Equipment Knowledge
- Professionalism

Expected time to complete scenario is 20 minutes.

Background/ Profile

Name: Jonny Dunn
Age: 50s
Works in the city as a banker
Personal Background

- Wife and 3 children.
- High blood pressure, penicillin allergy, no other medical issues
- Had tetanus many years ago

Settings

The scenario takes place in a simulated A&E cubicle. All equipment necessary to complete the scenario is available.

Persons present include:

4. Participant Trainee Doctor (being assessed)
5. Simulated A&E nurse
6. Simulated patient with traumatic wound/laceration

Scenario Sequence

Prior to the scenario, you were out drinking with colleagues following payday/bonus. You slip and catch yourself on a metal fence as you leave the taxi on the way home. You ask the taxi driver to take you to the nearest A+E.

You are sat on a chair with some gauze lightly covering your wound which the triage nurse applied. You are moderately intoxicated.

The nurse has performed a basic set of "obs" – the patient's blood pressure (115/70mmhg) pulse (68bpm) and oxygen saturations (99% on air), which are normal.

The participant doctor should:

- Examine the wound
- Ask some relevant history and tetanus
- Irrigate the wound
- Draw up and Apply local anaesthetic
- Open suture pack
- Suture wound

Throughout the scenario, the participant should be interacting in a professional manner with the nurse and patient.

On completion of suturing, the nurse will tell the doctor that she will sort out the dressings and that she can get on with the documentations. This will be the signal to end the scenario.

During the scenario you should:

- Ask about the "injection" and suture to see if it is necessary

- Ask the doctor if he/she is qualified/consultant
- Ask about taking out sutures
- Be a bit of a “moving target” – make it a little be challenging for the doctor to position your arm

Dress, Behaviour, Affect and Mannerisms of the Simulated Patient

- You are dressed in “work” clothes?
- You are moderately intoxicated with alcohol
- You are rude, abusive demanding and uncooperative.
    - Boast about your bonus and come out with some derogatory comments
    - Flirt with the nurse
    - Complain about the waiting time.
- You flinch when you see the needles and when the needle punctures your skin (until you are anaesthetised/ numbed)
- If it does not appear that the doctor put enough local anaesthetic then show pain when he/she attempts to close the wound

Appendix 4: Rating Tools and Assessment Criteria for Hybrid Simulations

A) Modified OSATS checklist score for wound closure*

Item	Not done/Incorrect	Done Correctly
1. Uses simple interrupted stitches for closure	**0**	**1**
2. Bites appropriate distance from skin edge (2-5mm)	**0**	**1**
3. Spaces sutures appropriately (2-5mm) and evenly	**0**	**1**
4. Skin edges approximated throughout incision	**0**	**1**
5. Skin edges everted with closure	**0**	**1**
6. Follows curve of needle- on greater than 80% of bites	**0**	**1**
7. Loads needle correctly ½ to 2/3 down needle- on greater than 80% of bites	**0**	**1**
8. Square knots with appropriate tension (no air knots)	**0**	**1**
9. At least 3-4 throws on nylon suture knots.	**0**	**1**
10. Appropriate handling of needle (ie. uses forceps)	**0**	**1**

*Modified from OSATS tool for excision of skin lesion and skin closure. Only items relating to skin closure were included in this version

B) OSATS Global Rating Score and Criteria

	1	2	3	4	5
Respect for tissue	Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time and motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument handling	Repeatedly makes tentative or awkward moves with instruments		Competent use of instruments although occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness
Knowledge of instruments	Frequently asked for the wrong instrument or used an inappropriate instrument		/knew names of most instruments and used appropriate instrument for task		Obviously familiar with the instruments required and their names
Use if assistants	Consistently placed assistants poorly or failed to use assistants		Good use of assistants most of the time		Strategically used assistant to best advantage at all times
Flow of procedure and forward planning	Frequently stopped operating or needed to discuss next move		Demonstrated ability for forward planning with steady progression of operative procedure		Obviously planned course of operation with effortless flow from one move to the next
Knowledge of specific procedure	Deficient knowledge. Needed specific instruction at most operative steps		Knew all important aspectsof the operation		Demonstrated familiarity with all aspects of the operation.

C) DOPs Rating Tool and Criteria*

Demonstrates understanding of indications, relevant anatomy, technique of procedure	**1 2** Mostly incorrect choice of management, instruments, anaesthetic and sutures	**3 4** Mostly correct choice of management, instruments, anaesthetic and sutures	**5 6** Completely correct choice of management, instruments, anaesthetic and sutures
Explanation of intervention including obtain patient consent to proceed	**1 2** Consent poorly obtained / No attempt No /poor explanation of procedure	**3 4** Consent Adequately obtained Satisfactory explanation of procedure	**5 6** Excellent consent Excellent explanation of procedure
Demonstrate appropriate preparation pre-procedure	**1 2** Inadequate preparation of equipment Inadequate patient apositioning	**3 4** Satisfactory preparation of equipment Adequate patient positioning	**5 6** Well organized, excellent preparation of equipment Good patient positioning
Appropriate analgesia/sedation / anasthesia	**1 2** No or poor use/administration of anaesthesia/analgesia	**3 4** Satisfactory use/administration of anaesthesia/ analgesia	**5 6** Excellent use /administration of anaesthesia/ analgesia
Technical Ability	**1 2** Poor technique Poor usage of instruments Poor economy of motion Lack of tissue respect	**3 4** Satisfactory technique Satisfactory usage of instruments Satisfactory economy of motion Satisfactory Lack of tissue respect	**5 6** Excellent technique Excellent usage of instruments Excellent economy of motion Excellent of tissue respect
Aseptic technique /clinical safety	**1 2** Poor aseptic technique Unsafe practice	**3 4** Satisfactory aseptic technique Generally safe practice	**5 6** Excellent aseptic technique Excellent safe practice
Post-procedural management and closure	**1 2** Inadequately explains to patient what happens next and closure	**3 4** Adequately explains to patient what happens next and closure	**5 6** Thoroughly explains to patient what happens next and closure
Seeks help where appropriate / team work/interaction	**1 2** Does not demonstrate ability to work with others	**3 4** Good teamworking and rapport with others.	**5 6** Excellent teamworking and rapport with others.
Communication skills, assessment of patients needs throughout scenario	**1 2** Poor ability to communicate Lack of empathy Use of jargon	**3 4** Good ability to communicate Demonstrates empathy Little use of jargon	**5 6** Excellent ability to communicate Demonstrates mpathy No use of jargon
Professionalism	**1 2** Lack of professionalism Lack of respect for patient Poor attitude	**3 4** Shows respect to patient Good attitude	**5 6** Clear professional approach Clear respect for patient Excellent attitude
Overall ability to perform the procedure	**1 2** Requires more training	**3 4** Competent	**5 6** Excellent

Appendix 5 – Post Simulation Questionnaire

Hybrid Simulation Study = Post Simulation Questionnaire

Clinician code: **Clinician Initials:**

Simulation date:

Simulation scenario: **1 □ 2 □**

PTT□ PFHS □ PFHS 2 □

Below are a number of statements regarding the simulation that you have just experienced. Please read each statement carefully and indicate your level of agreement or disagreement by circling the appropriate number on the scale below.

	Strongly Disagree				Strongly Agree
1. This scenario in the simulation approximates closely to what happens in the workplace	1	2	3	4	5
2. The wound is a realistic representation of a real wound	1	2	3	4	5
3. The equipment used in the simulation is realistic	1	2	3	4	5
4. The environment is a realistic representation of the workplace	1	2	3	4	5
5. I was fully immersed in the simulation	1	2	3	4	5
6. I performed as I do in the real workplace	1	2	3	4	5
7. I behaved in the same way as I do in the real workplace	1	2	3	4	5
8. The simulation allows me to adequately demonstrate my technical skills	1	2	3	4	5
9. The scenario allows me to adequately demonstrate my clinical knowledge for wound management and closure	1	2	3	4	5
10. The simulation allows me to adequately demonstrate my aseptic technique	1	2	3	4	5
11. The simulation allows me to adequately demonstrate my team working skills	1	2	3	4	5
12. The simulation allows me to adequately demonstrate my professionalism	1	2	3	4	5
13. The simulation allows me to adequately demonstrate my communication skills	1	2	3	4	5
14. The simulation is able to highlight strengths and weaknesses in my workplace performance	1	2	3	4	5
15. The simulation is an accurate judge of my overall competence in managing wounds	1	2	3	4	5

Appendix 6 Sample Interview for PTT vs PFHS Study (Chapter 5)

Interview for hybrid simulation study, candidate number AHM007, 30th June 2010. So thank you very much again for allowing, participating in this study. Would it be OK if I asked you a few questions about your simulations?

Yeah sure.

Just again to reiterate, the purpose of the study is to look at alternative forms of assessing doctors' competence in procedural skills. And the three scenarios you had, the first one was a part task trainer bench top scenario where you had a wound pad in front of you
And the second two scenarios were what we term hybrid simulation scenarios. Just really want to know what your initial thoughts are, how did you feel?

Yeah it was quite good, the first one was quite strange just because normally when you do just something with a clinical tool there's other people around, and you're not trying to do it in an *OSCE* style on, and go through things methodically, and it's actually quite hard to get into the mindset of treating it like a real wound, and also it just doesn't look desperately realistic. When you've got a real patient to interact with, I was quite surprised, I didn't really know what was going to happen in the simulation scenarios, but it was actually quite nice because you're having to manage talking to the patient, asking relevant things about the history, and then managing the wound itself, but it felt a lot more natural, and there were sort of the prompts of having the patient and the history in front of you made you think about elements of the wound management that were important.

So you said that it felt quite real. Do you mean real in comparison to the part task trainer, or do you mean in real terms of real world?

It was certainly real in terms of, very real compared to the past, part task trainer, but it was quite realistic having, the wounds themselves are fairly well positioned, and look like they are almost part of the limb, and at times I actually did forget that it wasn't the patient's genuine skin. And having the patient there to talk to, and also having the emergency nurse practitioner, made it feel much more like a real environment, and everything just about the set up, because you have the trolley and the drapes, and the couch for the patient, and all the rest of it, just makes it feel not exactly like, but very close to a kind of A&E scenario.

So it seems like there were quite a few elements that you described there that were realistic, were there any components in there that you thought distracted from the realism, or?

A couple of things where obviously because it's a simulated scenario I guess just like, well you know, with the aseptic technique, there are times when just using the normal gloves and things, well OK for the sake of the argument these are sterile, in real life you wouldn't do that. And just minor things like when injecting the Lidocaine because of the silicone it's worked straight back out the hole, which made me laugh, and wasn't terrible realistic. But most of it was actually not that, there weren't many distracting features certainly.

And do you think this overall realism affects your ability to perform?

I think it does but in a good way because if you are the best person in the world that's suturing a part task thing, and then you find one attached to a patient who's moving and can't deal with that side of things, then your skill isn't desperately useful. And also there are the more complex factors of the wound management, and how it was sustained, and just things like tetanus jabs and all the rest of it, which you're not really going to think about so much with the part task one, but you would want to think about with a real patient, and so I think in that sense the prompts are there and you're more likely to behave in a natural way, and if you were using it for assessment I think you'd get a much better idea of how candidates would behave in real life.

So you said that there's this more natural approach, do you think you were able to behave and perform naturally in that *arena*?

Yeah, as I say there were a couple of things where I possibly was not quite behaving as I normally would, for one thing I'm just not as experienced with suturing wounds, so I'd actually probably ask for help quite a lot more, or at least a level of supervision initially. And apart from that I think it was pretty realistic, the actors were very good, it was quite easy to be fairly natural with them, and Ewan was very realistic, I don't know if he is an ENP but it was

He's actually an A&E nurse I think, actually he is yes.

Yeah, so that was very kind of natural, and that really helped actually, having someone who is the, there as another member of staff, makes it feel more like real life. So yeah I think I did.

So do you think this is an acceptable way of assessing your clinical ability?

Yeah I think it is, I think there are areas as I say, where if I knew that I were being assessed, or if it were real life, I'd probably be a bit more stringent, but because possibly rushing a bit, and some of the prompts initially, like the sterile flasks, weren't there, and you just say, OK for the sake of argument these are sterile. But I think in general it is a very fair assessment because you have all the distracting features, and you have to just try and get on and do it, and just things like keeping an area sterile when it's on a patient is more difficult than keeping an area sterile when it's on a bench, and you can assess I think my ability to do that or not to do that appropriately, and exploring the wounds and all the rest of it. So yeah I thought it was a good tool for assessment.

Just a little point, which is not a question.

Yeah.

There were actually sterile gloves available, yeah.

Yeah, you only found them half way through the first one, but yeah.

But they were there, yeah

Yeah.

So I apologise for that.

No that's all right, that was

The next question I just want to ask you is relating to comparing those hybrid simulations versus part task trainer scenarios.

Respondent indicates agreement, yes

Do you think one or the other is a better or worse demonstration of your overall clinical competence, particularly in this case to wound management?

I think that the hybrid ones are still a better demonstration of my clinical competence because of the prompts that, about the history, how the wound happened, whether you need to look at foreign bodies, likelihood of there being a deep penetrating injury, and contamination, age of the wound, all of those factors, which are just much more natural to ask about in, when you've got someone to ask compared with when you've just got a piece of rubber and you have to try and list things off the top of your head, which is never a very realistic way of going about things, it's helpful but it's not as helpful I think.

Was there any difference in the way you behaved between the two different scenarios, different types of scenarios?

Yeah, I think when I was doing the part task one it was very kind of, OK, yeah, I think I'd do this, and I wasn't behaving as I would in a clinical environment, and I wasn't interacting with the piece of sponge in the way that I interact with a patient, no I definitely behaved much more naturally in the hybrid scenario.

Am I right in saying that, what you've just described in the last few things was in a part task trainer you're trying to show what you know by saying them out loud

Yeah.

But in the simulation you're just doing?

Yeah.

A more reasonable point there, yeah?

Yeah I think that's definitely the case, I think in, partly I wasn't really sure if I was meant to be talking through what I was doing, and I was partly talking through it, but it felt kind of in that like an OSCE style, fairly artificial scenario, whereas it, exactly in the hybrid you just get on with it and try and do the assessment management at the same time, and get the history one, both of those are going on.

And apart from what you just described earlier as the things that you felt were unrealistic about the part task trainer

Respondent indicates agreement, yes

You just said, OSCEs were quite unrealistic, in what way do you find them unrealistic?

I think the problem with the sort of traditional medical school OSCE is that for example, if you have to take a history for someone with jaundice you have to get a history of absolutely everything including whether they're a heavy drinker, whether they've had sex with any prostitutes and got risk factors for hepatitis, or any other … things that I would not get onto in five minutes talking to a patient in real life, because you need to build up a rapport first. Whereas in an OSCE when you've got a five minute station, then you rattle through things that quite honestly I just wouldn't even come onto in the first five minutes of talking to a patient, and I that I think it would be quite inappropriate to do. In more practical OSCE stations, some of them are close to what you'd do in real life, and some of them just, it is just trying to rattle off as many things as you can think of that are relevant. But when you've got a hybrid scenario, so for example the woman that had fallen off the bike, there were associated things to ask, it was important to make sure that the injury in the thigh wasn't a distracting injury, and that she didn't have anything else bad going on, she didn't have a head injury, the whole patient needs assessing not just a wound, whereas with the part task thing then you're just very focused on this small square of sponge that you have to put stitches in.

Did you find any difference in your motivation and attitude when you approached the two different types of tasks?

Yeah, when I was doing the part task one it was sort of, OK, right, fine, yeah, how do I do this? I can sort of remember, and just trying to talk myself through it. Whereas when I had a very realistic patient scenario it was more, OK, right, well at least the talking to the patient bit I can do, and then the prompts were more there. It was good actually having practised on the part tasks trainer first I think before doing the scenario, because I was slightly more familiar with the tools because I just don't suture very often, so to go from this for me was, it was quite helpful to do that. And then I think the difficulty with the hybrid one was for example, someone who's fairly inexperienced at suturing, like me, and wound management, if I'd have had the challenging patient as my first I think that would have quite possibly wrong footed me. If you're very experienced in that environment I don't think it would bother you as much, but for me it was quite helpful to go in the order I did.

We do actually swap the scenarios around to allow for negating the learning effect actually.

Respondent indicates agreement, yes

Now did you find the hybrid simulations more or less challenging that the part task trainers?

I actually felt that they were easier.

Good.

Just because they felt more natural, and because I like talking to people not a camera, so it was sort of yeah, and they were quite enjoyable actually.

To be honest you're not the first person to say that actually.

Yeah.

And I think that's, and again potentially one of our aims. Right, the last question I want to ask you is, the last few questions I want to ask you are relating to the two hybrid simulation scenarios. Did you find there was any difference between the two of them?

Just a bit yeah, you had one very pleasant, very coherent, very organised, calm lady who gave a good history of exactly what had happened, and remained completely still and helpful throughout, and then we had the man who'd had an unspecified amount of alcohol, and was very vague about how he incurred his wound, and was trying to get up every 30 seconds, and was being rude and abusive, and racist, and various other things at different points. So yes, one of them was very much more managing a wound and talking to the patient, ensuring there was nothing else, whereas the other one, there was a lot of patient management in it, and I think interestingly the things that I remember to ask, a couple of the things I remembered to ask the nice complaint patient, I forgot to ask the

Such as?

Such as drug allergies, I remember now, one of the things that again I don't think I would forget with any real patient, but again maybe a slight confounding factor, or maybe I would, and head injury, again, important, especially in someone who's behaving inappropriately. So it is important but because you get embroiled in the patient management side of things I guess that's when you probably … certainly might be that level, be wanting a team and support, and all the rest of it, and so, OK, I don't want to suture this guy while he's like this, can you get some more help? Because there were risk factors, and even just things like I had the arm cleaned and then, it was only at that point that I went, we need to get consent for this, because he was being very uncooperative, and yeah, it was just making sure that he was going to let me do it, and that it was going to be safe to do it. So yeah, I think it was, it was realistic, but I was certainly very different, and the fact that it was more challenging I think was reflecting in my performance and remembering certain things, so yeah.

Now it's been said more than once that a clinical environment, a real clinical environment, is, can be messy, chaotic

Respondent indicates agreement, yes

Do you think that these scenarios, the hybrid simulation scenarios, was in any way able to reflect that?

I suppose that most of the areas that have had to do any kind of, it's not really wound management, but anything vaguely clinical in, have been at least reasonably ordered, apart from chest drain insertions on wards, but yeah I think it was about the same in terms of the amount of space available was about what I'd expect in a bay in A&E, or, and it was I think as organised as you normally find things, I don't think it was particularly disordered, but then I don't think that most of the environments that I work in are that disordered anyway. …

Now have you got any questions for me?

Gosh probably. I'm just trying to think. Do you want to pause it while I …

It's all right, carry on.

So in terms of your application of this do you think you would be likely to use a combination of part task and hybrid scenarios to assess the same skill, or are you looking to more to just a hybrid, or does it depend on the results, or

Well at the moment it's still research *phase*

Sure.

But from a theoretical perspective it seems like quite an attractive solution to some of the limitations to just part task trainers, and obviously one of the things you've clearly highlighted was, people who are complete novices who need to get their basic … motor skills sorted out before they get thrown into that scenario

Yeah.

END OF INTERVIEW

Appendix 7 Participant Information Sheet and Consent Form

Imperial College London
100 years of living science
100

Participant Information Sheet
Basic Cardiovascular Examination Course – Ventriloscope® Validation

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?
Simulation is playing an increasing role in assessment of doctors' competence. We are investigating the validity of using a Ventriloscope® simulator to train medical students in basic cardiac examinations and auscultation.

Do I have to take part?
It is up to you whether or not to take part. If you decide to take part, you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason.

What will happen to me if I take part?

4. You will be asked to participate in a training course over two days and then assessed on a third day.
5. You will be asked to fill out questionnaires and answer questions in a focus group.
6. Your performance in the scenarios will be recorded and rated for data analysis.

What are the possible disadvantages and risks of taking part?
We do not anticipate any disadvantages.

What are the possible benefits to taking part?
The possible benefits of this study are to learn new clinical skills in basic examination and auscultation as well as increasing our understanding of the use of simulation for the purpose of training.

What if something goes wrong?
We do not foresee anything going wrong. However if there are any problems, you are free to terminate the simulation or post-simulation questionnaire/interviews at any time without giving any reason.

Will my taking part in this study be kept confidential?
All the data we collect from you will be anonymised by allocating you a code. Any paperwork containing your identity (consent form) will be kept separate to anonymised data collected during the study. They will not be shown to anyone outside the research team and can be destroyed as soon as your participation in the project comes to an end if you request.

What will happen to the results of the research study?
The results will be presented in departmental meetings, international conferences and peer reviewed journals. All data presented will be anonymous and never related to any specific individual or team.

Who is funding the research?
This research is funded by Imperial College.

Who has reviewed the study?
This study was given a favourable ethical opinion for conduct by St Mary's Research Ethics Committee.

Contact for further information
For more information, please contact:
Jimmy Kyaw Tun email: (Lead researcher)
Roger Kneebone email: r.kneebone@imperial.ac.uk (Research Supervisor)
Michelle Gatter email: michelle.gatter07@imperial.ac.uk (Researcher)

Additional researchers:
Monal Wadhera
Junaid Zaman

Imperial College
London

Participant Identification Number:_________________

Consent Form

Study Title: Basic Cardiac Examination Course – As part of the Ventriloscope® Validation Study

Principle Researchers: Dr. Jimmy Kyaw Tun, Dr. Roger Kneebone, Dr. Monal Wadhera, Michelle Gatter

PLEASE READ AND COMPLETE THE CONSENT FORM THOROUGHLY

- I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions ☐
- I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without legal rights being affected ☐
- I understand that my anonymised evaluations will be looked at by responsible individuals from our research team at Imperial College and that my data will be collated with other participants' data so that I cannot be identified individually ☐
- I understand that my performance in the simulations will be rated by responsible individuals and it does not impact upon my academic record ☐
- I agree to take part in the above study ☐

________________________ ______________________ ____/____/_____

Name of participant Signature Date

Appendix 8 Outline of Cardiovascular Examination Skills Course - Tutorial Slides

Objectives
General principles of clinical examination Patient interaction Learn Basics of CVS Examination Correlate Basic to Clinical Science Objective Structured Clinical Examination
Course Overview

Day 1
    Basic Science and Theoretical Underpinning
    Examination of the normal patient
Day 2
    Recap on examination
    Examination of patient with abnormal pathology
Day 3
    Formative Assessment and Feedback

Study overview

Purpose
    Explore use of Ventriloscope® to train CVS examination
    Compare training with Ventriloscopes to Real patients
Data Collection
    Demographics
    Pre- and Post- Questionnaires
    Performance rating
Consent

Overview of Cardiovascular Examination

End of bed inspection - 'look before you touch'
Hands – any clues to systemic disease?
Pulse and blood pressure – more than just numbers
Head – facies, eyes, mouth
Neck – arterial and venous waveforms
Chest –
    Inspection
    Palpation
    Auscultation + dynamic maneouvres
Abdomen – liver, spleen, AAA
Legs – oedema, cap refill, pulses.
Extra investigations

General Principles of Clinical Examinations

Patient Interaction
    Communication and professionalism
    Dignity and empathy
    Pain and comfort
    Asepsis and cleanliness
 	
General Principles of Clinical Examinations
Inspection
Percussion
Palpation
Auscultation

CVS Examination Overview
Video Introduction Set the Scene Introduce yourself

Check patient details
Explain you would like to examine the heart
Ask patient's permission to perform examination
Wash hands
Ask if patient is in pain
Expose patient appropriately
Position patient comfortably at 45°

Position for CVS Examination
Examine the patient from the right side at 45°
Peripheral Examination

End of bed inspection
Inspection of hands
Radial pulse
Blood pressure
Inspection of eyes, face and mouth
Jugular venous pressure (JVP)
Carotid pulse

End of Bed Inspection

Comfortable or distressed?
Short of breath or normal respiration?
Thin or overweight?
Obvious scars
Look around the bed
- Drugs
- Oxygen
- Cigarettes

Inspection of Hands

Clubbing
Peripheral cyanosis
Palmar erythema
Nicotine stains
Janeway lesions
Osler nodes
Splinter haemorrhages
Tendon xanthomas

Pulses

Radial
- Rate
- Rhythm

Brachial
- Blood pressure

Carotid
- Character
- Volume

Radial Pulse

Measure the *rate* and *rhythm*
On radial aspect of wrist ~2cm below the base of thumb
Always use your index and middle fingers to palpate the radial pulse
Do not use the thumb as it has its own pulsations
Count the rate for at least 15 seconds

Radial Pulse

Rate
- Normal sinus 60-100 bpm

Sinus bradycardia < 60 bpm
Sleep, fitness, drugs, hypothyroidism, heart block
Sinus tachycardia > 100 bpm
Exercise, pain, fever, drugs, hyperthyroidism, cardiac conduction disease e.g. atrial fibrillation

Rhythm
Regular = sinus rhythm
Regularly irregular = ectopic beats
Irregularly irregular = atrial fibrillation

Brachial Pulse and Blood Pressure

Measure the blood pressure (BP)
Pressure of blood in the arteries
Measured in millimetres of Mercury (mmHg)
BP is recorded in two figures e.g. 120/80 mmHg
Top number is systolic blood pressure
This is the pressure in the arteries when the heart contracts
Bottom number is systolic blood pressure
This is the pressure in the arteries when the heart rests between each heart beat

---

Blood Pressure

Normal BP = <140/90mmHg
Hypertension = >140/90mmHg
Essential hypertension
Drugs
Kidney disease
Endocrine disorders
Hypotension = <90/60mmHg
Sepsis
Hypovolaemia e.g. blood loss, burns, dehydration
Anaphylaxis
Post myocardial infarction

Face - Inspection

General
Anaemia
Malar flush

Eyes
Corneal arcus
Conjunctival pallor
Xanthelasma

Mouth
Cyanosis
Dental caries

Malar Flush
Corneal Arcus
Conjunctival Pallor
Xanthelasma
Central Cyanosis
Dental Caries

Jugular Venous Press
Correlates to right atrial pressure
Not palpable
Examine the *right internal jugular vein* (IJV)
Position patient at 45° with head turned slightly to the left
Locate the surface markings of the IJV
Runs from medial end of clavicle to the ear lobe

Between the two heads of sternocleidomastoid muscle Jugular Venous Pressure Jugular Venous Pressure Measuring vertical distance between the sternal angle and the top of the JVP JVP should be <3cm above the level of the sternal angle JVP raised in heart failure, fluid overload
Carotid Pulse: Palpation Palpate Palpate carotid pulse under angle of jaw Use light pressure only Never compress both at the same time Left thumb for right carotid and vice versa Comment on character and volume Precordium Examination

Appendix 9 – OSCE Rating Tool for Cardiovascular Examination

Introduction	Good	Adeq	Inadeq
Introduces self, states role and cleans hands	G □	A □	I □
Obtains consent and patient's details	G □	A □	I □
Correct exposure and patient positioning	G □	A □	I □
General Inspection			
Inspection from end of the bed (Well/ unwell, SOB, Oxygen, Nutritional Status)	G □	A □	I □
Peripheral Examination			
Inspection of Hands	G □	A □	I □
Palpates for Radial Pulse (Rate and Rhythm)	G □	A □	I □
Check for radio-radial delay		A □	I □
Asks for blood pressure		A □	I □
Inspection of face and oral cavity	G □	A □	I □
Examines JVP		A □	I □
Palpates for Carotid Pulse (Character)	G □	A □	I □
Precordium Examination			
Inspects chest for scars, visible palpations	G □	A □	I □
Palpates for apex beat		A □	I □
Palpates for heaves and thrills		A □	I □
Auscultates whilst palpating carotid for timing		A □	I □
Auscultates aortic area		A □	I □
Auscultates pulmonary area		A □	I □
Auscultates tricuspid area		A □	I □
Auscultates mitral area		A □	I □
Manoeuvre patient to the left and auscultates mitral area with bell		A □	I □
Manoeuvre patient to the left and auscultates left axilla area with bell		A □	I □
Lean patient forward and listen to left sternal edge for aortic regurgitation at end of expiration		A □	I □
Auscultate carotid artery for radiation of aortic stenosis		A □	I □
Auscultate lung bases		A □	I □
Student asks to perform other examinations (Sacral oedema, peripheral pulses, fundoscopy, urine dipstick)	G □	A □	I □
Closure			
Thank patient and cleaned hands	G □	A □	I □
Student ensured patient maintained dignity and pain free	G □	A □	I □
Student communicated with patient well throughout examination	G □	A □	I □
Diagnosis			
Summarises findings in a logical manner	G □	A □	I □
Provides diagnosis with reasoning	G □	A □	I □

Overall rating (GRS)	**1**	**2**	**3**	**4**	**5**	**6**
	Poor		**Satisfactory**		**Excellent**	

Appendix 10 Excerpt from Focus Group Transcription for Study in Chapter 6

Key Researcher

Before I start opening up up this discussion I just wanted to reiterate first to the study, so what we are looking at using stimulation. In this case in the form of Ventrilo scopes and simulated actor patients in the training of cardiac examination skills. We are going to focus on a few different questions regarding this study. Before we start I just want to get everyone to introduce themselves:

I am Jimmy Tun, Key Researcher on this project. Monal Wadhera, Key Researcher.
XXX first year medical student. Ventriloscope group
XXX, first year medical student Ventriloscope group.
XXX first year medical student Ventrilo scope group.
XXX, first year medical student, real patient group.
XXX first year medical student in the Ventrilo scope group.
XXX First year medical student in Real Patient group.

Key Researcher

First question I am going to ask is what do you guys think about the last few days in general, nothing specific, in general. How did you guys feel? Anyone want to start.

Medical Student
I thought it was a really good experience and I learned a lot, a lot of new skills and I thought it was really good.

Key Researcher
Anything a bit more specific. What was good about it?

Medical Student

Well I was in the Real Patient Group, so I thought it was really useful having a chance to listen to real patients and you know comparing to the use of multimedia as well.

Key Researcher
So how did you find listening to a real patient different from listening to the multimedia that we showed you - differs from the stuff from the internet?

Medical Studen
Of course, it is different in every patient. It is not always as obvious in a real patient as in the multimedia. I mean that is also something I would say about the ventriloscope as well. You know from what we have just done. I mean, because of the ventriloscope it makes it really obvious, like the sounds, but of course in a real patient it is not always going to be that obvious, and also it is like the ventriloscope is sort of operated by someone else , so you know they just see where you place the stethoscope. It might not necessarily have been in the right place for real patients, but it is just plays the sounds then.

Key Researcher

Okay anyone else have any thoughts about that sort of stuff?

Medical Student
I agree with that because I was on the ventriloscope group so when we are doing it, it was quite obvious to some, even if you lowered a sound, you still could hear it, whereas in a real patient, I found out today, between the real and the exam, with the real patient you had to position it right and you had to kind of put it sort of around the area just to make sure you eventually heard something or you haven't heard something –anything, so that it makes a big difference.

Medical Student
I thought the training, I think the initial ***block****. It was kind of usual, but I thought it may have been a bit too much at times. It may be me being a bit sleepy. But I just thought that there was a fair amount of a bit of information overload, especially in the beginning. Well eventually it all made sense when I was actually*

examining the patient, but to be fair I was getting kind of confused. I think the ventriloscope does have some slight issues in that it is bit too easy to hear and another issue is ***in simulated*** *patients. It is one of those* ***Simulated patients and real patients*** *but there are signs that way that you don't actually get to actually see* ***proper*** *signs and symptoms so when I went into the examination today I had no idea to* ***lesion*** *actually looked like in real life. I just saw them in the pictures.* ***Something*** *which looked a bit like a lesion/nodes but I wasn't sure.*

Key Researcher
That makes complete sense. I am just going to put this discussion back a little bit so in terms of the signs and symptoms and in terms of how easy and difficult it is, I guess that is partly to do with how we run the training. It was quite intensive the training course and that is why it was an intensive training course, actually so we appreciate it was probably a lot of hard work, particularly at your stage of training. In terms of whether we can do anything with a ventriloscope, we can actually have a judgment as to differing positioning of a patient, so if you don't position it correctly, you should not get the correct sounds. We can factor that into the design of the scenario of the patient encounter and also if there is anything that is slightly softer also, but though I agree with you certain signs we may not be able to replicate so at certain time we can, say for example, a mainline flush. So I guess that is what we are trying to do here is also tease out the limitations of how far we can use these equipments. In terms of when both yourself and ? Actually did the ventriloscope group here, how did you find it the first time you interacted with a real patient?

Medical Student
I didn't find it too difficult, because I think the first patient I was with he had quite obvious sounds so it was quite similar, but once I moved on to the other patients I found it more difficult because it was harder to get used to real patients. It is just not as simple as using the ventriloscope.

Key Researcher
What was difficult about it?

Medical Student

Well listening to the sounds, sometimes the murmurs themselves weren't as obvious and weren't as pronounced, so it was hard to work out where they were in a heartbeat and everything.

Key Researcher
Of course, working out heart sounds is just a small part of clinical examination and actually a lot of it actually you start without even touching the patient so what about the rest of the patient interaction? How did that feel?

Medical Student
Well, the rest of it was pretty much the same but when it came to the signs as ***Tigo s****aid, it was more difficult to recognize signs because I hadn't seen any before.*

Key Researcher
Where there any signs that you saw for the first time and went 'ah I didn't expect that'?

Medical Student
There was a big surgical scar on one of the real patients, I think actually I didn't actually notice that until I looked really close. The big white stroke which is quite obvious?

Key Researcher
Did anyone else get that?

Medical Student
No I didn't.
I think I noticed a pacemaker on one of them so that was something.

Medical Student
The scar really surprised me, because I noticed the pacemaker which was quite obvious especially when you move, but I didn't notice the scar because obviously he had hair on the chest which was what I expected

because I had gotten into such a routine it was very much a look at the chest but appears to be nothing obvious and ***didn't know how to look.*** *So I didn't notice.*

Medical Student
Because we had done it so much with people that looked perfectly normal and it is only that when you get to the listening, that we start really tuning and thinking there is going to be something. ***You ??*** *so that when we did* ***general action*** *it was very much do it quickly and get listening.*

Key Researcher
So let's say for example we repeat the ventriloscope and SP training, how would you change it to the correct this issue?

Medical Student
Personally I would basically use the ventriloscope as a starting off stage, so if it was in a compare research I would it use that to start off learning about it and then after that I would use real patients with actual symptoms and make them more realistic.

Medical Student

I think the ventriloscope is very good regarding what the sounds are like and also how they should vary in difference as well.

Key Researcher

I am going to concentrate on just you two now, XXX and XXX. In terms of when you guys encountered a ventriloscope for the first time today, how did that go?
Medical Student

Well, it felt like, the sounds you hear with any recording. It didn't feel that real.
Key Researcher
But what about interacting with a patient?

Medical Student
Interacting with a patient is similar really. But it does feel a lot more artificial. I noticed of course when you are doing a real patient ***assessment****, the sounds are not always as obvious as I said before. With the ventriloscope it is is really very obvious.*

Key Researcher
So perhaps the one thing we can do with a ventriloscope is to make the sounds less obvious which we actually can do actually.

So when you actually for the first time actually put a stethoscope to a real patient's chest, how confident were you?

Medical Student
I was fairly confident but basically when I turned the patient to the left I thought I heard a bit of murmur and I was expecting to hear radiation to the axillary. But it was a very faint sound but I wasn't sure whether it is possible to hear it sometimes just even though he or she doesn't have a murmur. It was too faint for me to actually make a decision there. I suppose it is due to lack of experience.

Key Researcher
But in terms of when I say interaction with the patient, I guess what I am really trying to get as, were you comfortable handling a patient?

Medical Student
Actually at times I found the real patients actually slightly easier to do than the ventriloscope. It might just only be the ventriloscope acting up because I did hear some very irregular rhythms with the ventriloscope.

Key Researcher

There were one or two technical issues at one of the stations actually, but in terms, none of you have ever (until the last few days) ever interacted with a real patient. What I really am trying get at is how comfortable were actually touching a patient for the first time?

Medical Student

I think it was a really accelerated course obviously. I thought it was brilliant. I actually really enjoyed this sort of close small teaching because if you have a big teaching group, ***you don't necessarily start off looking for stuff and that takes forever ….teaching group.*** *I actually thought the ventriloscope was good. I thought because when I was dealing with a real patient, I knew where to look for things and even this threw me off because this was with a real patient, I was still fairly confident that I was looking in the right place.*

Medical Student
Yes. It think it helps a lot that we had the practical directly after the lecture.

Key Researcher
What was the effect of having a practical straight after the lecture on the material inside the lecture?

Medical Student
It could understand it.

Key Researcher
Can I hear from the others can you tell me what did you think about the training course?

Medical Student
I thought it was really interesting to learn about the science bit and then interact with patients and actually listen to sounds. I thought it was pretty interesting.

Medical Student
I thought the course this year was overall pretty good, but generally at lectures I thought it was bit off and the ***lecturer*** *it seemed that it was throwing stuff at me that I had no idea about so that was slightly confusing. Everything got* **pointed to** *at that point kind of by the afternoon bit. I must say that those two first hours were probably my least favourite part of the course, overall I enjoyed it.*

Medical Student
I personally thought there were a bit of information overload, like ***XXX s****aid, but I realized that it is just part of the process for me kind of start with and looking at the bigger picture so it is a lot to learn obviously, but looking at the bigger picture and actually consolidate down to how you can put this into practice and yeah, I really enjoyed it.*

Medical Student

I thought it might have better if it was broken up slightly, so maybe if it was an hour of basic skills or an hour of skills, like physiology, then be a practical and then going into more, say the pathology of certainties, especially of the hand, I found hand diseases slightly confusing, especially all you could tell was a few figures on a screen and no real.

Medical Student
I really enjoyed the first bit of the first two hours. I think it was taught very well and in a very accessible way. I think it might just be my learning style is different possibly and I kind of get along with that kind of teaching a bit better. I thought the amount of information we got was the right level, quite accessible and I thought because we had the practical soon afterwards, we could put into context and I would seeing things I remembered from the lecture and looking for things that I remember from the lecture when we doing the examinations. I think it has helped to connect everything quite well.

Key Researcher

Okay brilliant. Just following on from what Jimmy said, if I can ask the ventriloscope participants, he was asking about if you felt confident enough to then go and see a real patient. Did you think it was a useful

exercise or to use the ventriloscope before you see a real patient or would you had been happy to have gone and examined a real patient anyway?

Medical Student

I think that starting with ventriloscope helps a lot to be honest, because really it is difficult to find patient with that kind of heart problems first and along with that what I think the essential factor of getting used to this kind of examination is actually just getting used to the routines and getting your skills fluent. To be honest, we actually had an actor acting out to be a patient is only a matter of the sound using the stethoscope, so when I actually went in to see a patient it was just like that. No different at all.

Key Researcher

What do the other guys think?

Medical Student

*Well, yeah I think the ventriloscope is really good because I think it is really important to have a very basic understanding and know the sounds, even if they aren't really clear, then you know after that you are listening for certain sounds and so at times there would be a variation in different sounds, other sounds coming in as well which you do get with the real patient. I do think you do feel a lot more confident once you start with the ventriloscope and essentially is more accessible because you can change the sound, you can make different sounds and it's focusing on a single symptom as opposed to the patient. When I **walked** in to the patient what threw me off was that there were so many different things. There was the scars, the pacemaker, looking at his hands and his condition wasn't as good as a normal patient. So it was nice having the ventriloscope to have a really good focus on just listening to heart sounds which could vary.*

Medical Student

Made easier instead of having quite a few patients and bringing them in with different conditions that you to sort through. Things like that.

Key Researcher

So you are suggesting that it might be easier to use the ventriloscope for maybe some then

Medical Student

Especially for this much accelerated kind course. I think it's much easier for the people organizing it as opposed to getting all sorts of different patients in order to keep within hours? You just have more choice then.

Key Researcher

Medical Student

I really liked using the ventriloscope because I think you had to hear the heart sounds for different moments at some point and with the multimedia it was very clear and good, but it was completely out of context listening to it on over speakers and when you do it with the ventriloscope you have the stethoscope in different positions, used the bell on the diaphragm and you hear the sounds more in context to where they would be. I thought it was quite helpful.

Appendix 11 Sample Anonymised Free-Text Response to Knowledge and Understanding Questionnaire

VS Group – Question Item 2

Q2 What are the key components of a cardiac examination?

Participant	Pre-Course	Post-Course
VS1	Listening to the patient's chest and heart	Inspection – chest, face, peripheral region Palpation – for heaves and thrills, pulses Auscultation – to aortic pulmonary, tricuspid
VS2	Informing the patient of what is about to happen. Listening to rhythm/ value sounds	Communication Inspection Palpation Auscultation
VS3	Listening for heart sounds, checking pulses, looking at hands	Examining for peripheral signs – hands, face (especially) feeling pulses. Palpating chest Auscultating to listen for abnormal sounds
VS4	Looking, feeling, listening and stethoscoping	Palpation Inspection Auscultation
VS 5	Viewing patient/ looking at general demeanour. Listening to heart sounds, history taking, ECG	Inspection Palpation Percussion Auscultation
VS6	Speed of beat, character of sound (valves)	Noticing signs and symptoms (peripheral and central)

Appendix 12 - JKT Unstructured Observation Account of Cardiovascular Examination Training Course Session

This is a study looking at the teaching and learning of basic cardiovascular examinations comparing the use of real patients with a simulated-patient/part-task simulator (ventriloscope) hybrid simulation. My aim at this point was to observe how learning (or lack of learning) took place in general in both conditions. I was particularly interested in differences in learning as well as how simulation is utilised. Other more general observations are made throughout to help shed light on other qualitative (focus group and pre and post-training comments) and quantitative data (performance of students) collected.
My role in the course was purely overseeing and managing the course, therefore freeing myself to make observations of both the teacher and students. Aside from being involved in the design of the training course, I was predominantly a non-participant observer.
The aims of the skills course was to teach medical students basic cardiovascular examination skills. The level of the teaching was designed to match what a 3rd year medical student would be expected to learn (1st year of clinical training).

There were two courses running in parallel, one with patients with real clinical signs who were invited from the cardiology outpatient clinic in the usual manner one would for OSCEs and one with hybrid simulation and ventriloscopes. These were two days in total followed by a final OSCE assessment comprising of 4 cardiovascular examination stations: 2 RP 2VS

	RP group	VS Group
Day 1	Lead by Tutor JZ Interactive tutorial (didactic component) Hands-on practice of basic cardiovascular examination on normal subject	None
Day 2	None	Lead by Tutor JZ Interactive tutorial (didactic component) Hands-on practice of basic cardiovascular examination on normal subject
Day 3	Lead by Tutor MG Review of principles of clinical examination Hands-on practice on two different real patients with different clinical signs None	
Day 4		Led by Tutor MG Review of principles of clinical examination Hands-on practice on two different simulate patients/ventriloscope hybrids with different clinical signs
Day 5	OSCE with real patients and simulated patient/ventriloscope hybrid	OSCE with real patients and simulated patient/ventriloscope hybrid

Day 1: RP Patient Group Account
Present

– Students A-F. These are 1st and 2nd year medical students at Imperial College, who are in their "pre-clinical" years of training and have had no prior experience in cardiovascular examination.
- Tutor JZ is a year 4 cardiology registrar, who has had extensive experience in teaching and examining in medical student assessments (OSCES).
- JKT Key Researcher
- MG Key Researcher (my research assistant)/ BSc dissertation student

The day started with introduction of the course by the tutors. Students gather around most of them looking quite enthusiastic about the prospect of learning something clinical. For this cohort of students, although they had enrolled to study medicine, they have had little patient / real patient contact. They are perhaps not yet what Lave and Wenger would describe as legitimate peripheral participant in relation to being a clinician. Whilst most students appeared enthusiastic, I believed not all of them may be what is thought of as classical "adult learners". There was also a mix of student types and not all students were equally enthused. Student B was what I would describe as every tutors "ideal" student – enthusiastic, knowledgeable and receptive. This variation is a key issue when researching educational activities (which I explore in more detail later.) Nevertheless, the course got of to a good start, starting with a series of interactive tutorials. These tutorials comprised of an overview of the key components of the basic cardiovascular examination. Tutorials were delivered using conventional teaching aids, i.e. PowerPoint slides etc. For clinical signs, particularly

It was interesting watching the students attempt to learn, what was usually taught predominately in the clinical years. Traditionally, UK medical courses were divided into: pre-clinical component, where all basic sciences aimed to prepare the student for clinical learning; and a clinical component, where the main emphasis is learning through being in the clinical environment; though its worth bearing in mind that in recent years, there has been a move towards greater integration of the two.

I was uncertain at this point as to whether my students would be able to learn the material provided to them given the relative lack of pathology knowledge and understanding – indeed there were times throughout these tutorials where certain terminologies seem unfamiliar to the students. A number of students asked what infective endocarditis was, whereas others queried the various valvular pathologies. Knowledge of anatomy and physiology were however generally good, which is unsurprising given that the students would have recently completed modules in these topics as a result of the year of training. I am however pleasantly surprised by how much some of the students know – I don't recall knowing very much as a first year student.

The tutorials seemed to have a mixed reception, some students being more engaged than others. It may be due to the difference in learner types within this group. Tutorials are in general within the comfort zone of these students, whereas, patient contact, at least initially appears not to be which I will go into more detail later.

In the afternoon, all the students were given the opportunity to learn how to perform the cardiovascular examination on a normal volunteer. Initially, tutor J proceed to demonstrate expertly how a "textbook" cardiovascular examination is performed. The examination was performed with a high degree of fluency and fluidity, whilst a certain air of professionalism was portrayed.

Interestingly, his demonstration was a performance – in the sense that he is being watched by a cohort of observers, but at the same time, as he goes through the various aspects of the examination, he makes statements of his intentions and findings,

> To patient, "Have you got any pain in your arm sir? May I raise it up to feel your pulse?"
>
> To students, "I am just checking for a water-hammer pulse"

In a series of repeat "performances" Tutor J makes a series of statements relating to the need to be observant and understand pathology

"Remember when we discussed the crescendo-decresendo nature of the pulse this morning…. (hand action) this is why you may feel a water-hammer pulse"

Throughout this process, the tutor states and reminds students the importance of professionalism and good patient care demonstrating his approach to these areas whilst examining the patients.

He also describes examples of real clinical situations – the relevant clinical context –

"So when I am examine a patient in clinic, sometimes I combine the examination of several sytems together (to increase efficiency)"

"It is important to get the patient in a good position, but this is not always possible, and you will have to learn how to make the best of your situation"

"I know everything seems quite rushed, especially when it comes to doing your OSCEs, but in reality, say when I am in clinic, you have a lot more time."

At the same time, tutor J provides a constant reminder of the other key objective, i.e. the need to pass examinations.

"You must wash your hands and introduce yourself … these are vital marks on the OSCE"
" Remember to ask for an ECG at the end of the examination"
The effect of this demonstration is evident. Students for the first time are seeing theory put to clinical practice. At the same time, there is an element of role modelling that seems to take place – students can see what they are meant to achieved, where they are meant to be after years of training. The tutor is painting mental images of what it is like to perform the examination "for real" on a patient in the real clinical setting.
What follows is a series of hands-on practice sessions where the students practice the sequence of the cardiovascular examination under the expert guidance of Tutor JZ. Initially, unsurprisingly there is an element of cognitive overload. The ability to integrate newly acquired knowledge and develop examination skills simultaneously is challenging to all the student.
It is interesting to see the tutor again balancing carefully the need to intervene and correct students with "letting them go" to experiment. This is a dilemma in much of medical training, and perhaps more so with respect to invasive procedures. Throughout the course of the session, what I observe is a gradual reduction in the tutor's input, particularly with respect to the routine of the examination. Not infrequently, the tutor will correct the students' techniques, some of which are to make the examination easier, such as by improving patient positioning, whilst at other times it is to help students master difficult techniques such as feeling for the apex beat*.

Student C spends about half a minute feeling for the apex beat. Tutor JZ recognises that student C is having difficulty finding the patients' apex beat. After locating the apex beat himself Tutor JZ guides student C's hand to where the beat is. Recognising that this can be tricky for other students, he asks everyone to gather round and "have a go" feeling for the apex beat.
There are also several occasions where the tutor reassures students, reminding them that they are learning a skill that takes years to master (I am also reminded of this, and am therefore mindful of what the outcomes may be when the students are tested).
It is interesting how much can be learnt from practicing on a normal human subject. In essence, it allows simulation of the routine of a procedure/ examination as well as allows students to identify what normal is.

Day 2 RP Group
Present
– Students A-F – Preclinical students as above
- Tutor MW is a Qualified GP and GP Trainer, who has had extensive experience in teaching and examining in medical student assessments (OSCES) as well having written exam preparation books for GPs.
- JKT Key Researcher
- MG Key Researcher (my research assistant)/ BSc dissertation student
- Volunteer Patient A
- Volunteer Patient B
Equipment and Setting: Clinical Skills Laboratory.

The morning session started off with a recap and overview of the cardiovascular examination, the tutor going over the salient features of a cardiovascular examination. The tutor then introduces the volunteer patient to the students reminding them to be respectful and mindful of the patient's needs at all times. We specifically asked the students to dress formally as they would on the ward.
The tutor then demonstrates the cardiovascular examination on the real patient to the team again showing the professionalism and respect most good clinicians would demonstrate.
Our volunteer patient (A) is in his mid 60's with aortic stenosis. Aside from being slightly less agile, he is or average to large build and fully co-operative. The examination is performed in the setting of a clinical skills

training laboratory and as a result, some of the challenging factors that may be present in a busy clinical ward are not represented in this setting.
Students then take turns examining the patient's cardiovascular system as the tutor guides them through the process. Perhaps as result of the presence of the tutor as well as learning the routine of the cardiovascular examination the previous day, the students do not seem to have any issues coming into contact with the patient physically, given that none of them had examined patients previously.

What is evident though is that some of the knowledge and skills gained on the previous day of the course have been forgotten, which is unsurprising given the students have only had 1 day of training and had previously no knowledge or the skills in this domain. They are however more familiar with the new language and terms required when learning to perform examinations. The range of terminology used, ranging from the clinical signs to specific manoeuvres can be quite alien. Rather than talking purely about endocarditis (inflammation, usually secondary to infection of the endocardium), students are now trying to familiarise themselves with terms such as, " Janeway lesions", " Osler's Nodes" and "Splinter haemorrhages. Whereas maneovres such as checking for the "apex beat" may sound more familiar, others such as "water-hammer pulse" which refers change in pulse-wave due to aortic regurgitation (detected by palpation) are perhaps not.

The student's examination are not fluent in many aspects. Most students are again trying to find the apex beat, which is in quite a different position to when they practiced on normal subjects. They also spend more time auscultating for the known cardiac murmur, trying to familiarise themselves with not just the sounds, but also the positions/ areas where the sounds are heard, e.g. radiation to the carotids. The theory that was learnt in yesterdays tutorials and multimedia files seem to become more consolidated through this process.
At one point, as only one student could examine the patient at anyone time, some of the students broke off and practiced on each other to rehearse the routine again and again, knowing that they will be tested the following day. The tutor also encouraged this.
The process of rehearsal is however conducted within a shorted time-frame, with aspects of the examination skimmed through. Less attention is payed to the "simpler" elements and where signs cannot be elicited. The process appears to be similar to memorizing a list,
Student C states, "I'm looking for splinter haemorrhages, Janeway lesions, Osler's nodes and pallor", then quickly moves to measuring the pulse rate. The time spent on the hands signs were minimal, though the process of measuring pulse rate was almost what would be expected in a real clinical encounter. My interpretation is that as the normal subject (fellow students) have no hand signs, it is easy skim over them, whilst even without an abnormal pulse, it is still possible for students to properly practice timing it.
The afternoon session is similar to the morning session in that the tutor provides a recap of the examination, followed by a new volunteer patient (B) in his early 60's who has Mitral Regurgitation and a different cardiac murmur.
The students are now more familiar with the routine of performing the cardiovascular examination and what is expected of them. This time the students are somewhat more fluent, though again when it comes to auscultation, students in general slow down, trying hard to pay attention to the subtleties of the sound, and practicing manoeuvres such as detecting axillary radiation of the sound.
Once again, I am noticing that the students perhaps spending less time on certain aspects of the examination, whilst focusing on areas that are either more challenging, or where a pathological sign is present.
By mid afternoon, it is evident that the patient is quite tired from have 6 medical students examining him repetitively. Both my research assistant and the tutor are aware of this and suggest a break.
There seems to be a degree of fatigue amongst the students as well - the students have repetitively practicing the routine for the greater part of the training day.

Day 1 VS Group
Present
– Students G-L. 1st and 2nd year "Preclinical" medical students at Imperial College, and have had no prior experience in cardiovascular examination.
- Tutor JZ (As above
- JKT Key Researcher
- MG Key Researcher (my research assistant)/ BSc dissertation student
Equipment and Setting: Clinical Skills Laboratory.

Similar start to the day as with RP Group starting with introduction of the course by the tutors. Students as with RP Group are enthusiastic about learning something clinical as they have had little clinical experience or patient contact.

There was again a mix of learner types in the group some seemingly more self motivated than others whilst others somewhat less enthused. Student I came in 1 hour late to the tutorial and only as a result of my research student chasing him up to attend. Instead of apologising, the student was somewhat blasé about the incident – this is reflected in the focus group data.

Tutor JZ again delivered the same interactive tutorials to the ones for RP students, which comprised of an overview of the key components of the basic cardiovascular examination. Tutorials were delivered using conventional teaching aids, i.e. PowerPoint slides etc. For clinical signs, particularly

It was interesting to see at this point, the students who have relatively little clinical knowledge to make the jump from learning basic science to learning about not only pathology, which they have some idea about, but also names for signs and examination manoeuvres. This represented a completely new language for many of the students. Names such as water-hammer pulse, Jane way lesions may seem alien to the students. Again knowledge of anatomy and physiology were however generally good, which is unsurprising given that the students would have recently completed modules in these topics as a result of the year of training. I am however pleasantly surprised by how much some of the students know – I don't recall knowing very much as a first year student.
The tutorials seemed to have a mixed reception, some students being more engaged than others. It may be due to the difference in learner types within this group. Tutorials are in general within the comfort zone of these students, whereas, patient contact, at least initially appears not to be which I will go into more detail later.

The afternoon session is designed to be predominantely practical whereby, all the students were given the opportunity to learn how to perform the cardiovascular examination on a normal volunteer. Tutor JZ demonstrates expertly how a cardiovascular examination is performed with fluency, confidence and professionalism.

As with when Tutor JZ demonstrates and guides the students to perform the cardiovascular examination, there are many incidences where his position is more than a teacher teaching technical aspects of the patient examination but as a role model, exhibiting a high level of professionalism. Although everyone knows that the "Patient" is in fact normal and a healthy volunteer, they all follow by (the tutor's) example and conduct themselves with a high level of professionalism and behave as if the volunteer is a real patient.

He again also describes examples of real clinical situations, and paints the clinical context as with when teaching the RP group.

Students then take turns examining the patient under the close guidance of the tutor. Each examination takes longer than would usually as expected, as each run of the cardiovascular examination is interrupted with questions and demonstration on finer aspects of technique such as for when looking at the JVP.

There are again multiple occasions where the tutor reassures students, reminding them that they are learning a skill that takes years to master.

Again, my feelings are that much of the routine in this example can be and is learnt on real patients.

Day 2 VS Group
Present
– Students G-L. 1st and 2nd year "Preclinical" medical students at Imperial College, and have had no prior experience in cardiovascular examination.
- Tutor MW (As above)
- JKT Key Researcher
- MG Key Researcher (my research assistant)/ BSc dissertation student

Equipment and Setting: Clinical Skills Laboratory.
Ventriloscope

The second day of the course is lead by tutor B (MW)
Following a quick introduction, the tutor provided an outline of the day's learning objectives was presented followed by a quick (30 minute) recap on the theory and principles of the cardiovascular examination.
Tutor B then demonstrates the Ventriloscope to the students to familiarise them with the device. The students had already familiarised themselves with the real stethoscope when practicing on normal subjects (i.e. each other) on the day before, the Ventriloscope is slightly different physically and in terms of the sounds played back, so we thought it was important to allow for familiarisation of this device.
The tutor then introduces the SP (Actor patient) and demonstrates a full cardiovascular examination on the SP using the ventriloscope.
Again as with in on the first day of the course, tutor B conducts the cardiovascular examination with a high level of professionalism and respect as most good clinicians would demonstrate.
The SP is in his 40-50's who has been asked to portray the role of a real patient unless asked otherwise. The SP was also asked to portray a patient who was fully co-operative, but only to comply with instructions for manoeuvres if they in the role of the patient understood the students. It is worth noting that these SPs including the ones used in the study generally have a vast amount of experience in terms of portraying the patient for many clinical encounters and are therefore very knowledgeable with respect to the process and content of the cardiovascular examination. However, in this course, they were asked specifically not to demonstrate this knowledge and play the role of a lay person.
To match the content of training course with real patients, the ventriloscope was set to simulate aortic stenosis for this session. The sounds were activated by tutor MW when the stethoscope was correctly positioned.
The students then took turns performing a cardiovascular examination on the patient. Although the students had a full afternoon practicing their examinations skills on each other the day before, there were many aspects of the examination, which were not fluent. This is unsurprising given that the students have only had one day of training in a skill which has multiple steps and components and can take a lifetime to master.
Interestingly, they did not seem to be phased by the having to perform on an SP. This is probably multifactorial. First the environment and situation was "safe". The SP portrayed a compliant patient and did not offer some of the challenges that may present in the real clinical work. Reflecting on my own experience, it is far harder to perform a good quality cardiovascular examination on a demented patient than it is on a young healthy patient attending a routine pre-assessment clinic. In addition, there was the presence of an "expert" tutor who is accessible for guidance and advice. Finally, although the students have not performed a cardiovascular examination previously, they have had contact and dialogue with real patients and SP previously as part of their medical curriculum.
During the process of the students performing the cardiovascular examination, as with in the previous day, the tutor intervenes at the request of the student or if she detects that the student is struggling. During this time, the tutor will often given the student reassurance and facilitate learning through demonstration. As observed in the previous day, one of the components of the cardiovascular examination that the students seem to struggle with is in trying to locate the position of the apex beat by palpation. Other examples of areas that the students appeared to find difficulty with are in palpating for the carotid pulse, identifying the jugular venous pulse waveform (JVP) and auscultating for cardiac sounds. When these difficult situations are identified the tutor often interrupts the student performing the cardiovascular examination and gathers the students to all get involved taking turns practicing the component of the cardiovascular examination that has been identified to be difficult.
During the session, through dialogue between the tutor and the students, the tutor would present some advice based on their own clinical experience. The tutor through direct instruction and role modelling demonstrated the type of professional behaviour and conduct when interacting with a patient. At the same time, tutor B like with tutor A from day one would stress the importance of thoroughness, ensuring all the steps of the cardiovascular examination are conducted. At times, it appears that the thoroughness is for good clinical practice, however, there are again certain signs that what is taught is for good OSCE "examination technique". For example, tutor B would stay to the students,

"…(Remember) to turn the head only slightly, otherwise it will be difficult to see the JVP"

"Introduction and washing your hands are easy points…. And you can fail the exam if you do not do these"

"Always remember to thank the patient"

Throughout the session, the students practice on the SP but also break away an practice on one another. Their fluency also develops and the amount of input required by the tutor diminishes.

In the afternoon session, the tutor again recaps on the principles of cardiovascular examination. A new SP is introduced and the Ventriloscope is now programmed to simulated mitral regurgitation. Again a similar process occurs whereby the tutor guides them through the cardiovascular examination, though this time, the students are generally more fluent. The area the appears to require the most guidance is in the auscultation, which is unsurprising given that it is the main difference between this session's simulation and the simulation used in the morning.

About midway through the session, a number of students breakaway from the main activity of examining on an SP and started practicing in other ways. Some students practiced palpating for apex beats and carotid pulses on one another, others concentrated on auscultation. Interestingly, with respect to the auscultation, there a number of students chose to practice on a table instead of on the SP, miming the placement of the ventriloscope on the table as if it was the patient. It appears, that although the students were provided with what I believed to be a holistic simulation, at times they preferred to focus on some key areas, often without the use of the SP.

Another phenomena that I witnessed when observing the students practicing on one another was to quickly skim through the process of the cardiovascular examination, completing what is usually an 8-10minute examining within 4-5 minutes. Even components that take time such as checking for pulse rate were skimmed over. It appeared that students were deliberately practicing in a shorted space of time for reasons of efficiency. My interpretation of this is that students by this point were quite confident with checking the pulse and so they just needed to "go through the motions". Perhaps a contributing factor to this was the fact that they knew they were going to be assessed, which was made clear by us as the researchers and reinforced by the teacher.

At the end of the session, the students appear to be quite exhausted, perhaps due to the intensity of the training and the amount of repetition. The tutor presents some final remarks and offers the students an opportunity to ask some final questions