

Assessment of Technical Skills In Surgery



**Doctorate of Philosophy
Imperial College
University of London**

**By
Sudip K. Sarker
Clinical Safety Research Unit
Imperial College London**

Supervisor

Professor Charles Vincent



Co-Supervisor

Professor Sir Ara Darzi

1st February 2007

Abstract

Aims

Assessing surgical technical skills in a structured manner is a topical issue in light of re-validation, shorter higher surgical training periods and high profile medico-legal cases. To date there has been little attempt to comprehensively assess both generic and operation specific technical skills in live open and laparoscopic surgery. We therefore aim to develop and validate new tools which can assesses both generic and specific technical skills in open and laparoscopic surgery. Also we aim to determine if there is a learning effect of novices learning to assess technical skills as well as to determine what type of surgical experience is required to reliably assess technical skills in surgery.

Methods

Previous assessment tools OSATS (likert) and the core skills in the JCHST (checklist) tool were adapted and changed to assess generic technical skills. For specific technical skills previous assessments tools and hierarchical task analysis were used to construct individual specific technical skills assessment tools for each operation. Three key operations were chosen, elective primary inguinal hernia repair, sapheno-femoral junction ligation and laparoscopic cholecystectomy. Full length versions of the operations were recorded onto DVD. Two surgeons with more than 12 years postgraduate surgical experience assessed each operation blindly and independently. All of the operation was viewed, and only non relevant parts were fast forwarded.

Results

195 operations were analyzed, 100 live laparoscopic cholecystectomies, 48 live primary inguinal hernias and 47 (live and simulated) sapheno-femoral junction ligations. These

were performed by a total of 17 Consultant Surgeons and 34 Registrar Trainee Surgeons. Reliability between the observers and construct validity between the two surgeons groups were demonstrated for generic and specific technical skills in laparoscopic cholecystectomies and inguinal hernia repair using the likert scoring. Reliability between the observers was good for likert scoring of generic and specific technical skills in live sapheno-femoral junction ligation, however construct validity was not demonstrated for generic and specific technical skills using likert scoring between the Consultant Surgeons and Senior Registrar Surgeons. Combining the live and simulated sapheno-femoral junction ligations demonstrated construct validity for generic technical skills between the Consultant Surgeons and all the Registrar Surgeons (Junior and Senior). Regarding assessing technical skills, there is a learning effect by novices who improve with training in assessing the technical skills in laparoscopic cholecystectomies. Also only a surgeon with an adequate amount of surgical and technical experience can independently and reliably assess the generic and specific technical skills in laparoscopic cholecystectomies.

Conclusions

This likert scoring assessment tool of generic and specific technical skills in open (inguinal hernia) and laparoscopic (cholecystectomy) surgery seems to have good inter-rater reliability. It seems to have face, content, concurrent and construct validities regarding generic and operation specific technical skills. Novices cannot reliably assess generic and specific laparoscopic technical skills, however they do improve with training. A fairly experienced surgeon can only reliably assess the technical skills in laparoscopic cholecystectomies. The assessment tools may have the possibility of being used in surgical training, education, appraisal and self-appraisal.

Table of Contents

Abstract	3-4
Acknowledgements	18
Declaration of Originality	19
Introductory Outline of Thesis	20
Diagram 1.0 - Structure of Thesis	21

Introduction

Chapter 1

Surgery, Outcome & Training

1.1 Surgery

1.10 Overview	22
1.11 Qualities of a Surgeon	22
1.12 History	23
1.13 Surgical Specialities	26
1.14 Surgical Pathway	28
1.15 Evolution in Surgery	29

1.2 Surgical Outcome

1.20 Overview	30
1.21 Outcome of Common General Surgical Operations	30
1.22 Adverse Events in Surgery	35
1.23 Determinants of Surgical Outcome	37
1.24 Conclusions	45

List of Tables

Table 1.20 Surgical Outcome Prognostic Scoring Systems	38
--	-----------

Table of Contents

1.3 Training in Surgery

1.30	Overview	46
1.31	History	46
1.32	Present Day	47
1.33	Modes of Training	48
1.34	The Future	52

Chapter 2

Assessment of Technical Skills in Surgery

2.10	Overview	53
2.11	Background	53
2.12	Method of Review	56
2.13	Current Methods of Assessment of Technical Skills	58
2.14	Structured Assessment of Technical Skills	60
2.15	Conclusions	71

List of Tables

Table 2.1 Principles of an Assessment Tool	54
Table 2.2 Current Methods of Assessment of Technical Skills	55
Table 2.3 New Methods of Assessment of Technical Skills	55
Table 2.5 Operative Competence JCHST	60
Table 2.6 Objective Structured Assessment of Technical Skills	62
Table 2.7 Likert & Checklist Scoring Papers	63
Table 2.8 Dexterity Analysis Systems Papers	66
Table 2.9 Virtual Reality Simulators Papers	68
Table 2.10 Error Scoring Papers	70

List of Pictures & Figures

Picture 2.1 Imperial College Surgical Assessment Device	73
Picture 2.2 ICSAD Tracing	73
Figure 2.1 Assessment of Technical Skill in Surgical Training	74

Table of Contents

Chapter 3

Human Error Reliability

3.10	Overview	75
3.11	Human Error – Definitions, Types & Classifications	79
3.12	Human Error Identification Assessment Techniques	86
3.13	Assessing Errors in Healthcare	89
3.14	Assessing Errors in Surgery	90
3.15	Conclusions	98

List of Tables

Table 3.0 Errors	77
Table 3.1 Error: Surgical Classifications	86
Table 3.2 Range & Scope of HRA Techniques	89
Table 3.3 Task Analysis of Laparoscopic Tissue Dissection	96
Table 3.4 Task Analysis of Laparoscopic Gallbladder Dissection	96

List of Figures

Figure 3.1 Rasmussen's Error Level Model & Reason's Error Type Classification	82
Figure 3.2 Reason's Model of Latent Failures	84
Figure 3.3 Terms Used in Task Analysis	92
Figure 3.4 Hierarchy of Creation of Pneumoperitoneum	97

Statement of Objectives of Thesis	99
--	-----------

Methodological Study

Chapter 4

Hierarchical Task Analysis in Surgery

4.10	Overview	101
4.11	Rationale of Surgical Task Analysis	101
4.12	Process of Constructing a Hierarchical Task Analysis	103
4.13	Phase 1 - Review Process & Planning	104
4.14	Phase 2 - Constructing Initial Templates of Task Analysis	105

Table of Contents

4.15	Phase 3 - Constructing Surgical Task Analysis	108
4.16	Phase 4 - Constructing an Anaesthetic Task Analysis	108
4.17	Phase 5 - Constructing a Scrub Nurse Task Analysis	109
4.18	Phase 6 - Constructing Surgical Hierarchical Task Analysis	110
4.19	Phase 7 - Refining Surgical Hierarchical Task Analysis	110
4.20	Phase 8 - Constructing Final Hierarchical Task Analysis	115
4.21	Conclusions	124

List of Diagrams

Diagram 4.1 Constructing a Surgical Hierarchical Task Analysis for an Operation	102
Diagram 4.2 Phases of Constructing a Surgical HTA for an Operation	103

List of Tables

Table 4.1 Initial Template of Task Analysis of a Laparoscopic Cholecystectomy	107
Table 4.2 Task Analysis of One Task in a Laparoscopic Cholecystectomy	108
Table 4.3 Task Analysis of One Task Demonstrating Surgeon & Anaesthetist Interaction in a Laparoscopic Cholecystectomy	109
Table 4.4 Task Analysis One Task Demonstrating Surgeon & Scrub Nurse Interaction in a Laparoscopic Cholecystectomy	110
Table 4.5 Checklist Scoring for Inter-Rater Reliability Tasks & Sub-Tasks	112
Table 4.1A - Total Scores for each Surgeon for Sub-Tasks – Inguinal Hernia	113
Table 4.2A - Total Scores for each Surgeon for Sub-Tasks – Sapheno-Femoral Junction	114
Table 4.3A - Total Scores for each Surgeon for Sub-Tasks – Lap Chole	115
Table 4.10 HTA of Primary Inguinal Hernia Repair Surgical View	116
Table 4.11 HTA of Sapheno-Femoral Junction Ligation Surgical View	118
Table 4.12 HTA of Laparoscopic Cholecystectomy Surgical View	121

List of Figures

Figure 4.1 Inter-Rater Reliability Sub-Tasks of Primary Inguinal Hernia Repair	113
Figure 4.2 Inter-Rater Reliability Sapheno-Femoral Junction Ligation	114
Figure 4.3 Inter-Rater Reliability Sub-Tasks of Laparoscopic Cholecystectomy	115

Table of Contents

Methodology

Chapter 5

Developing Criteria to Assess Generic & Specific Technical Skills in Surgery

5.10	Overview	125
5.11	Methods - Generic Technical Skills	125
5.12	Methods - Specific Technical Skills	129
5.13	Refinement & Feasibility of Scoring Systems	135
5.14	Reliability of Scoring Systems	136
5.15	Conclusions	137
	List of Tables	
	Table 5.1A Generic Technical Skills – OSATS Likert Scoring Assessment Tool	126
	Table 5.1B Generic Technical Skills – PhD Likert Scoring Assessment Tool	127
	Table 5.2A Generic Core Skills – JCHST Assessment Tool	128
	Table 5.2B Generic Technical Skills – PhD Checklist Scoring Assessment Tool	129
	Table 5.3 Creation of Pneumoperitoneum	130
	Table 5.4 Retraction & Exposure of Calot’s Triangle	130
	Table 5.5 Cystic Artery & Duct	130
	Table 5.6 Dissection of Gallbladder	130
	Table 5.7 Removal of Gallbladder	130
	Table 5.8 Specific Likert Technical Skills Assessment Scale	131
	Table 5.9 Eubanks Error Scoring – Laparoscopic Cholecystectomy	132
	Table 5.10 Seymour’s Error Scoring – Laparoscopic Cholecystectomy	133
	Table 5.11 Tang’s Error Scoring – Laparoscopic Cholecystectomy	134
	Table 5.12 Specific Technical Skills Checklist Scoring Minor Events	135
	Table 5.13 Specific Technical Skills Checklist Scoring Major Events	135
	Table 5.14 Specific Technical Skills Checklist Scoring Significant Major Events	135
	Table 5.15 – Individual Mean Scores & Inter-Rater Reliabilities Between Two Experienced Surgeons Laparoscopic Cholecystectomy	136
	Table 5.16 – Individual Mean Scores & Inter-Rater Reliabilities Between Two Experienced Surgeons Open Inguinal Hernia Repair	137
	Table 5.17 – Individual Mean Scores & Inter-Rater Reliabilities Between Two Experienced Surgeons Sapheno-Femoral Junction Ligation	137

Table of Contents

Chapter 6

Design of Studies

6.10	Overview	138
6.11	Design	138
6.12	Profile of Observing Assessors	138
6.13	Profile of Surgeons Assessed	139
6.14	Statistics Used & Definitions	139
6.15	Outline of Study Chapter Templates	139

Assessment Study

Chapter 7

Assessment of Technical Skills in

Open Live Inguinal Hernia Surgery

7.10	Overview	140
7.11	Material & Methods	140
7.12	Results	144
7.13	Conclusions	156
	List of Tables	
	Table 7.1A Generic Technical Skills Likert Scale for Inguinal Hernia Repair	158
	Table 7.1B Specific Technical Skills Likert Scale for Inguinal Hernia Repair	158
	Table 7.2A Generic Technical Skills Checklists Scale for Inguinal Hernia Repair	159
	Table 7.2B Specific Technical Skills Checklists Scale for Inguinal Hernia Repair	159
	Table 7.3A Means & Ranges of Likert Scores of Surgeons – Open Hernia	145
	Table 7.3B Means & Ranges of Checklist Scores of Surgeons – Open Hernia	145
	Table 7.4 Reliability of Technical Skills Checklists	151
	Table 7.5 Validity of Technical Skills Checklists	154

Table of Contents

List of Figures

Figure 7.1 Operation Grade & Level of Surgeon	147
Figure 7.2 Operative Time & Level of Surgeon	148
Figure 7.3 Generic Likert Technical Skills Reliability	149
Figure 7.4 Specific Likert Technical Skills Reliability	150
Figure 7.5 Generic Technical Skills Likert Scoring	152
Figure 7.6 Specific Technical Skills Likert Scoring	153
Figure 7.7 Generic Technical Checklist Scores	155
Figure 7.8 Specific Minor Technical Checklist Scores	155
Figure 7.9 Specific Technical Skills Checklist Major & Significant Major Scores	156

Assessment Study

Chapter 8

Assessment of Technical Skills in Open Live & Simulated Sapheno-Femoral Surgery

8.10 Overview	160
8.11 Material & Methods	160
8.12 Results	165
8.13 Conclusions	183

List of Tables

Table 8.1A Generic Likert Scale for Sapheno-Femoral Junction Ligation	185
Table 8.1B Specific Likert Scale for Sapheno-Femoral Junction Ligation	185
Table 8.2A Generic Checklists for Sapheno-Femoral Junction Ligation	186
Table 8.2B Specific Checklists for Sapheno-Femoral Junction Ligation	186
Table 8.3A Means & Ranges of Likert Scores Live Open SF	166
Table 8.3B Means & Ranges of Likert Scores Simulated Open SF	166
Table 8.4A Means & Ranges of Checklist Scores Live Open SF	167
Table 8.4B Means & Ranges of Checklist Scores Simulated Open SF	167

Table of Contents

List of Tables

Table 8.5 Reliability of Technical Skills Checklists Live Open SF	173
Table 8.6 Construct Validity of Checklists Technical Skills Scoring Live Open SF	175
Table 8.7 Validity of Technical Skills Checklists Live & Simulated Open SF Operations	180

List of Figures

Figure 8.1 Operation Grade & Level of Surgeon Live Open SF	169
Figure 8.2 Operative Time & Level of Surgeon Live Open SF	170
Figure 8.3 Generic Likert Technical Skills Reliability Live Open SF	171
Figure 8.4 Specific Likert Technical Skills Reliability Live Open SF	172
Figure 8.5 Generic Likert Technical Skills Scoring Live Open SF	174
Figure 8.6 Specific Likert Technical Skills Scoring Live Open SF	174
Figure 8.7 Generic Checklist Technical Skills Scoring Live Open SF	176
Figure 8.8 Specific Minor Checklist Technical Skills Scoring Live Open SF	176
Figure 8.9 Specific Major & Significant Major Checklist Live Open SF	177
Figure 8.10 Generic Likert Technical Skills Scores Live & Simulated Open SF	179
Figure 8.11 Specific Likert Technical Skills Scores Live & Simulated Open SF	179
Figure 8.12 Generic Checklist Technical Skills Scores Live & Simulated Open SF	181
Figure 8.13 Specific Minor Checklist Scores Live & Simulated Open SF	181
Figure 8.14 Specific Major & Significant Major Checklist Scores Live & Simulated Open SF	182

Assessment Study

Chapter 9

Assessment of Technical Skills in Laparoscopic Cholecystectomy Surgery

9.10 Overview	187
9.11 Materials & Methods	187
9.12 Results	191
9.13 Conclusions	210

Table of Contents

List of Tables

Table 9.1A Generic Likert Scale for Laparoscopic Cholecystectomy	213
Table 9.1B Specific Likert Scale for Laparoscopic Cholecystectomy	213
Table 9.2A Generic Checklists for Laparoscopic Cholecystectomy	214
Table 9.2B Specific Checklists for Laparoscopic Cholecystectomy	214

List of Tables

Table 9.3A Means & Ranges of Likert Scores of Surgeons	192
Table 9.3B Means & Ranges of Checklist Scores of Surgeons	192
Table 9.4 Reliability of Technical Skills Checklists Between Observers	200
Table 9.5 Validity of Generic Technical Skills Checklists	205
Table 9.6 Validity of Specific Technical Skills Checklists	205

Chapter 9

List of Figures

Figure 9.1 Gallbladder Grade & Level of Surgeon	194
Figure 9.2 Gallbladder Grade & Consultant Surgeon Years	195
Figure 9.3 Operative Time & Level of Surgeon	196
Figure 9.4 Gallbladder Grade & Operative Time of Consultant Surgeons	196
Figure 9.5 Operative Time & Consultant Surgeon Years	197
Figure 9.6 Generic Technical Skills Likert Scoring Reliability Between Observers	198
Figure 9.7 Specific Technical Skills Likert Scoring Reliability Between Observers	199
Figure 9.8 Generic Technical Likert Scores of Consultants & Trainees	201
Figure 9.9 Generic Technical Likert Scores & Individual Consultant Surgeons	202
Figure 9.10 Generic Technical Likert Scores & Consultant Surgeon Years	202
Figure 9.11 Specific Technical Likert Scores of Consultants & Trainees	203
Figure 9.12 Specific Technical Likert Scores & Individual Consultant Surgeons	203
Figure 9.13 Specific Technical Likert Scores & Consultant Surgeon Years	204
Figure 9.14 Generic Technical Checklist Scores of Consultants & Trainees	206
Figure 9.15 Generic Technical Checklist Scores & Individual Consultant Surgeons	206
Figure 9.16 Generic Technical Checklist Scores & Consultant Surgeon Years	207
Figure 9.17 Specific Minor Technical Skills Checklist Scores of Consultants & Trainees	207
Figure 9.18 Minor Technical Error Checklist Scores & Individual Consultant Surgeons	208

Table of Contents

List of Figures

Figure 9.19 Specific Minor Technical Skills Checklist Scores & Consultant Years	208
Figure 9.20 Specific Technical Skills Checklist Major & Significant Major Scores of Consultants & Trainees	209
Figure 9.21 Major Technical Checklist Scores & Individual Consultant Surgeons	209
Figure 9.22 Major Technical Skills Checklist Scores & Consultant Years	210

Assessment Study

Chapter 10

Assessors of Technical Skills in Laparoscopic Cholecystectomy Surgery

10.10 Overview	215
10.11 Methods – Study 10.1	217
10.12 Statistical Analysis	219
10.13 Results	220
10.14 Methods – Study 10.2	224
10.15 Statistical Analysis	225
10.16 Results	226
10.17 Conclusions	228

List of Tables

Table 10.1 BSc Students & Experienced Surgeons Mean Scores	220
Table 10.2 Percentage Agreement of BSc Students to Experienced Surgeons	223
Table 10.5 MSc Students & Experienced Surgeons Mean Scores	226
Table 10.6 Percentage Agreement of MSc Students to Experienced Surgeons	228

Table of Contents

List of Tables

Table 10.10 BSc Students (Before Training) & Experienced Surgeons Technical Skills Assessment Scores	221
Table 10.20 BSc Students (After Training) & Experienced Surgeons Technical Skills Assessment Scores	222
Table 10.30 MSc Students & Experienced Surgeons Technical Skills Assessment Scores	227

Conclusions & Applications

Chapter 11

11.10 Overview	230
11.11 Findings of the Studies	231
11.12 Limitations of Findings of Studies	232
11.13 Research	235
11.14 Implementations for Clinical Practice	238
11.15 Education	240
11.16 Summary	243

List of Figures

Figure 11.1 Basic Surgical Trainee - Curriculum of Assessment & Feedback	244
Figure 11.2 Higher Surgical Trainee - Curriculum of Assessment & Feedback	245
Figure 11.3 Basic Surgical Trainee - No Curriculum of Assessment & Feedback	246
Figure 11.4 Higher Surgical Trainee - No Curriculum of Assessment & Feedback	246

<u>References</u>	247-279
--------------------------	---------

<u>Publications & Presentations</u>	280-282
--	---------

Papers & Articles

Presentations & Abstract Publications

Table of Contents

Appendix

Appendix Chapter 1

Surgical Outcome

List of Tables

Table 1.1 Measures of Surgical Outcome	283
Table 1.2 Types of Operation Affecting Outcome	283
Table 1.3 Patient Factors Affecting Surgical Outcome	283
Table 1.4 Non-Patient Factors Affecting Outcome	284
Table 1.5 Organisation & Equipment Affecting Surgical Outcome	284
Table 1.6 APACHE II Score	284
Table 1.7 POSSUM Scoring	285
Table 1.8 Framework of Factors Influencing Clinical Practice	286

Appendix Chapter 2

Assessment of Technical Skills in Surgery

List of Tables

Table 2.4 Papers on Technical Skills Assessment 1975 to 2005	287-311
--	---------

Appendix Chapter 3

Human Error & Human Reliability 312

List of Tables

Table 3.5 – Surgical Task & Error Analysis Papers 1995 – 2005	314
---	-----

Table of Contents

Appendix Chapter 4

Hierarchical Task Analysis in Surgery

List of Tables

Table 4.13 Task Analysis of Inguinal Hernia Anaesthetist 's View	315
Table 4.14 Task Analysis of Inguinal Hernia Scrub Nurse's View	317
Table 4.15 Task Analysis of Long Saphenous Varicose Veins Anaesthetist 's View	318
Table 4.16 Task Analysis of Sapheno-Femoral Junction Ligation Scrub Nurse View	321
Table 4.17 Task Analysis of Laparoscopic Cholecystectomy Anaesthetist 's View	322
Table 4.18 Task Analysis of Laparoscopic Cholecystectomy Scrub Nurse's View	325

Appendix Chapter 7

Assessment of Technical Skills in Open Live Inguinal Hernia Surgery

List of Tables

Appendix 7.1	326
Appendix 7.2	326
Appendix 7.3	326

Acknowledgements

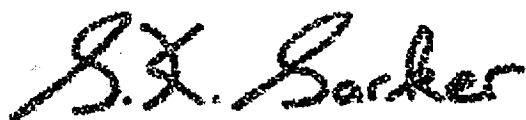
I am thankful for the help from the Academic Surgical Unit staff at St Mary's Hospital, in particular to the surgical staff, my fellow research colleagues and secretarial staff. I am also grateful to the clinical medical and research students who participated in some of the studies, and members of the Clinical Safety Research Unit at St Mary's for their advice and comments on this research project. I am indebted to the help, advice and mentoring I have received by my supervisors Professor Charles Vincent and Professor Sir Ara Darzi who have both guided me throughout this research project. Charles has been inspirational and has been the ultimate Academic whilst I have been writing my PhD thesis, always giving me helpful and constructive advice. Furthermore, I would like to thank Ms Avril Chang, Senior Lecturer in Surgery, without her participation and help this whole project could not come to fruition.

This thesis covers the field of assessing technical skills in surgery. I had the great fortune whilst writing up my thesis to work for Mr Adam Lewis, Consultant General and Colorectal Surgeon at the Royal Free Hospital, one of the greatest colorectal surgeons I have seen operate. He depicts personally and professionally everything a surgeon should be: intellect, compassion, natural technical ability, an educator and above all, a gentleman. He has inspired me and nurtured my own abilities during the time I worked for him.

Finally, I thank Charlotte my wife, her family and my own family for encouraging me through this project. They have all given me tremendous support and praise during this period in my surgical life. I dedicate my thesis to my father who sadly died after I completed my thesis.

Declaration of Originality

I, Sudip Kumar Sarker declare that the contents of this Doctor of Philosophy thesis submitted to Imperial College of Science, Technology & Medicine, University of London has been originally researched and written by myself.

A handwritten signature in black ink, appearing to read 'S.K. Sarker', written in a cursive style.

1st February 2007

Signature

Date

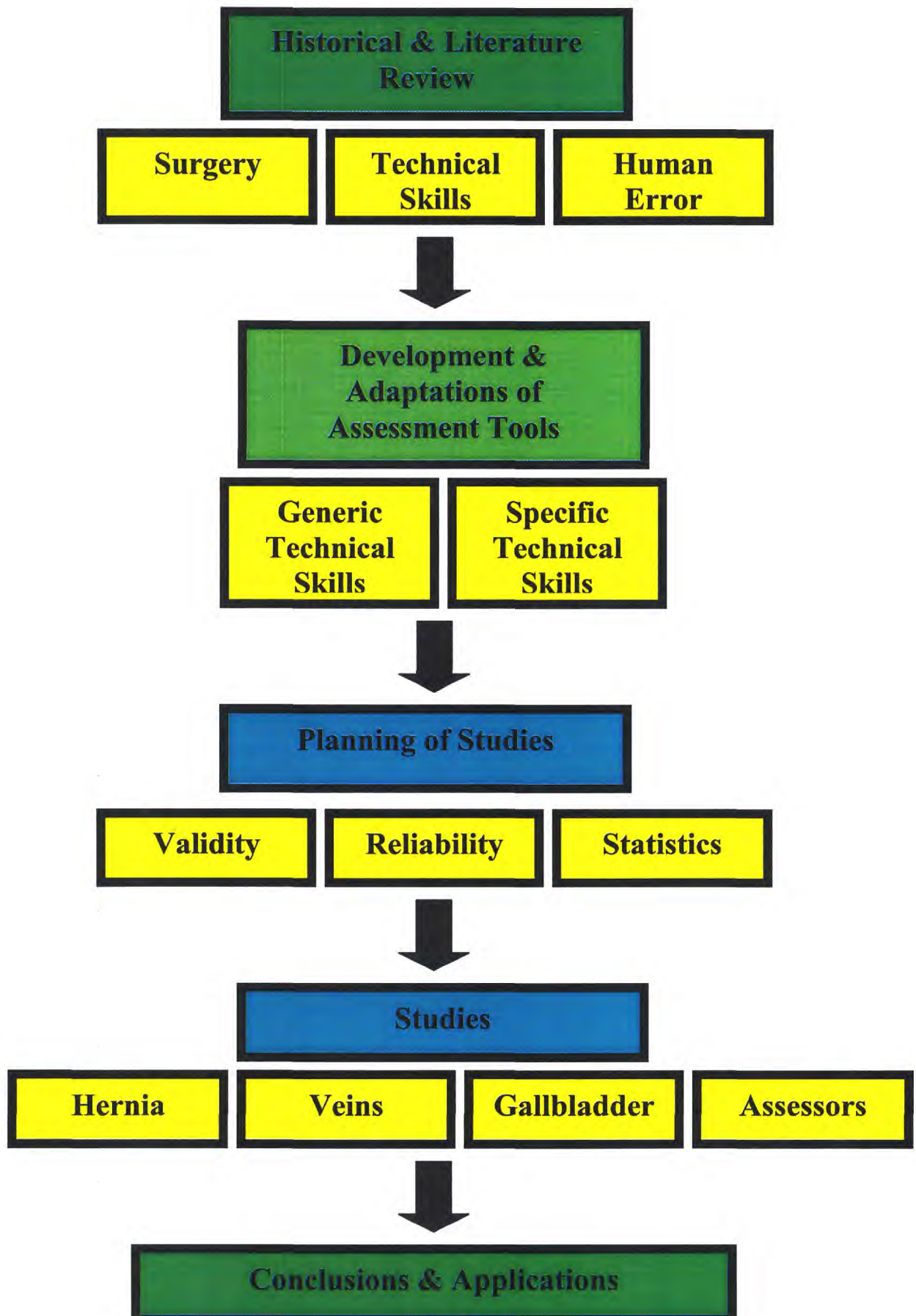
Introductory Outline of Thesis

Acquisition of technical skills in surgery has been passed down through each generation over the last few thousand years. Traditionally these skills are learnt in an apprentice style way of learning. After seeing several operations, the trainee surgeon is then supervised by his or her mentor to perform the operation under supervision. Then after the trainee surgeon was deemed competent he or she is then allowed to operate solo. However there is no structured assessment of the trainees' technical ability and competency.

In the 21st century the acquirement and competency of technical skills by surgeons is a topical issue. The accountability of surgeons' technical ability and how they learn these skills has come into close scrutiny. There is no widespread assessment tool to assess the technical skills of surgeons performing simple to complex open and laparoscopic surgery, assessing generic technical skills as well as the quality of the operation, which is reflected in specific technical skills.

In this thesis we aim to develop tools to assess the generic and specific technical skills of expert and trainee surgeons performing open and laparoscopic in live and simulated surgery. In developing these tools we have used previous assessments tools as references and used hierarchical task analysis in developing these assessment tools. We aim to demonstrate their reliability; face, content, concurrent, predictive and construct validities. In the later part of the thesis we assess the different types of surgical experience which are required to reliably assess technical skills of surgeons when performing laparoscopic surgery. Finally there is a discussion of how the assessment tools and its methodologies can be applied to current surgical training, appraisal, education and clinical applications. The structure of the thesis is depicted in the flow diagram 1.0.

Diagram 1.0 – Structure of Thesis



Chapter 1

Surgery & Training

1.1 Surgery

1.10 Overview

Surgery is concerned with the treatment of diseases, deformities, and injuries using manual procedures which require dexterity. Surgery can be used to repair broken bones, stop uncontrolled bleeding, and remove injured or diseased tissue and organs. People have practiced Surgery since ancient times, but it did not become a respected science until the 19th century. Increasing knowledge of the human body, the discovery of anaesthesia, and the use of sterile operating procedures combined to make Surgery a safe and effective method of treatment. In the 20th and 21st century advances in technology have helped the field of Surgery grow at a rapid pace, with new techniques ever growing. This section summaries the qualities of a surgeon, history of surgery, surgical specialities, surgical pathway and evolution in surgery.

1.11 Qualities of a Surgeon

The Roman Celsus suggested that a surgeon should be youthful, have a strong and steady hand, ambidextrous, have a clear vision and not lose concentration by the cries of the patient. Thomas Vicery in 1633 said to be a perfect surgeon one must be firstly learned, secondly an expert, thirdly ingenious and finally well mannered. With the invention of anaesthesia operations did not have to be done hastily. Lister lectured to his students that with the advent of anaesthesia the need for operative speed was abolished and allowed time for care in the procedure, adding success depends on attention to detail. Greenberg in 1982 asked a group of surgeons which personality traits were appropriate for a surgeon. The top

traits were decisiveness, fairness, good team participation, flexibility, honesty about errors, discipline, the ability to consider all the facts, motivation and the ability to listen.

It is therefore important to assess the right properties of a good and competent surgeon. Surgical competency is being proficient in surgical technical skills, team performance and decision making. This thesis will mainly concentrate on the structured assessment of surgical technical skills.

1.12 History

The first ever surgical procedures were performed in the Neolithic Age (about 10,000 to 6000 BC). Burr holes, a procedure in which a hole is drilled in the skull to relieve pressure on the brain, may have been performed as early as 8000 BC. In Egypt, carvings dating to 2500 BC describe surgical circumcisions.

Ancient Egyptian medical texts have been found that provide instructions for many surgical procedures including repairing a broken bone and mending a serious wound. In ancient India, the Hindus surgically treated bone fractures and removed bladder stones, tumours, and infected tonsils. They are also credited with having developed plastic surgery as early as 2000 BC in response to the punishment of cutting off a person's nose or ears for certain criminal offences. Using skin flaps from the forehead, Hindu surgeons shaped new noses and ears for the punished criminals. In the 4th century BC, the Greek physician Hippocrates published descriptions of various surgical procedures, such as the treatment of fractures and skull injuries, with directions for the proper placement of the surgeon's hands during these operations.

During most of the Middle Ages (5th century to 14th century AD), the practice of surgery declined. It was viewed as inferior to medicine, and its practice was left to barbers who

travelled from town to town cutting hair, removing tumours, pulling teeth, stitching wounds, and bloodletting.

In 1316 the French surgeon Guy de Chauliac published *Chirurgia Magna* (Great Surgery). This massive text describes how to remove growths, repair hernias, and treat fractures using slings and weights. The text helped surgery gain respect as a creditable science. At this time a new order of surgeons arose in France. They were called surgeons of the long robe, distinguished from the barber surgeons who were known as surgeons of the short robe. The barber surgeons had little medical training, while the surgeons of the long robe were studied physicians and considered such practices as bloodletting primitive.

During the 16th, 17th, and 18th centuries, many discoveries in surgical practice took place. Much credit belongs to the French surgeon Ambroise Paré, often called the father of modern surgery. Paré successfully employed the method of ligating, or tying off, arteries to control bleeding. Discoveries about functions of the human body also helped make surgery a more accurate science during this period. For instance, the English physician and anatomist William Harvey discovered the process of blood circulation and Italian anatomist Marcello Malpighi identified the existence of tiny blood vessels called capillaries that carry blood from the major blood vessels to the cells of the body. John Hunter, a British anatomist and surgeon, stressed the close relationship between medicine and surgery and performed many experimental operations that advanced the practice of clinical surgery.

Most surgery, however, continued to be restricted to less critical areas of the body or to operations that did not penetrate the skin too deeply. Surgeons rarely opened the abdomen, chest, or skull because of the pain it caused the patient and the risk of infection. This changed in 1846 when anaesthesia was used as a way to mask pain during surgery by American dentist William Morton. Although Morton is often credited with the discovery of

surgical anaesthesia, American surgeon Crawford W. Long used anaesthesia in 1842 during the removal of tumours but did not publish his results until 1849.

Post-surgical infections remained a serious complication of surgery until the mid-19th century when the French chemist Louis Pasteur discovered that fermentation or putrefaction, the decay and death of body tissue, is caused by bacteria in the air. In 1865 the British surgeon Joseph Lister applied Pasteur's work to surgery, developing antiseptic (germ-killing) techniques including the use of a carbolic acid spray to kill germs in the operating room before surgery in Glasgow. These antiseptic procedures helped eliminate postoperative infection. Other physicians, including Austrian Ignaz Semmelweis and American Oliver Wendell Holmes, determined that bacteria are also carried on the hands and clothing and transferred from patient to patient as a physician attends one after another. These physicians pioneered techniques such as washing hands and changing into clean clothing before surgery that prevented wounds from being contaminated during surgery.

In the late 1800s, having solved the problems of pain and infection, surgeons began performing new types of surgery including procedures on the abdomen, brain, and spinal cord. At the turn of the 20th century, improved diagnostic abilities and methods of treatment helped surgery to become even more effective. When the German physicist Wilhelm Conrad Roentgen invented X rays in 1895 to "photograph" the inside of the body he changed the way surgery was performed. The discovery of the blood groups A, B, and O by Austrian pathologist Karl Landsteiner enabled surgeons to give patients transfusions of their own blood type to ensure survival during and after surgery.

Other technological advances permitted surgeons to perform increasingly complex and difficult operations. The introduction of antibiotics in the 1940s further minimized the risk of postoperative infection. The development of the heart-lung machine in 1953 by

American surgeon John H. Gibbon allowed surgeons to more easily and successfully perform surgery on these organs. It also marked the beginning of modern heart surgery. The operating microscope, developed in the 1950s, provided surgeons with a way to perform delicate operations on minute body structures like the inner ear and the eye, and more recently, enabled surgeons to reattach the tiny blood vessels from severed limbs to the body. The first kidney transplants were performed in the 1950s, and the first heart transplant, in 1967, was performed by South African surgeon Christiaan Barnard.

1.13 Surgical Specialities

Surgery is performed by specially trained doctors known as surgeons. General Surgery and the Surgical Specialties training is conducted in association with a hospital and usually lasts from six to ten years depending in which country the training is performed. At the end of this period, known as a residency or higher surgical training, the surgeon may receive further training to learn the skills of a particular specialty, or subdivision, of surgery. The following is a simplified overview of General Surgery and the Surgical Specialities and demonstrates the wide range of technical skills required by a surgeon.

General Surgery is the broadest surgical division, focusing on surgery of the abdomen, the breast, and the endocrine organs. In addition in the U.K. some general surgeons also perform vascular surgery, and those Consultant Surgeons who do not have a cardio-thoracic department in their hospital, may have to operate on the heart and lungs in the trauma and emergency situation.

Paediatric Surgery focuses on the unique conditions of operating on infants and children

Colon and Rectal Surgery procedures are performed on the anus, rectum, and intestines.

Neurological Surgery involves operations on the brain and spinal column. Also head injuries are dealt by neurosurgeons, and emergency burr holes and craniotomies may need to be performed.

Breast Surgery in the U.K. is performed by general surgeons specialising in this. Some general surgeons completely specialise in this field and perform reconstructive breast procedures.

Gynaecology encompasses a variety of procedures, including surgery to remove diseased reproductive organs such as the uterus and procedures to correct female infertility or facilitate permanent contraception.

Obstetrics is a division that focuses on all aspects of a woman's pregnancy and may involve procedures such as a caesarean section.

Ophthalmic Surgery involves operations on the eye and often requires the use of microsurgical techniques performed under a microscope.

Orthopaedic Surgery involves operations on bones, muscles, and joints as a result of disease or trauma.

Otolaryngology involves the surgical treatment of diseases of the ears, nose, tongue, larynx (vocal cords), and neck.

Plastic Surgery encompasses cosmetic procedures to improve appearance and reconstruct damaged parts of the body such as skin and underlying muscle.

Cardio-Thoracic Surgery deals with surgery of the lungs, chest wall, heart, and large blood vessels of the chest.

Vascular Surgery involves replacing or repairing blood vessels, particularly arteries that deliver oxygenated blood to the body tissues. In the U.K. general surgeons sub-specialise in this, unlike in the U.S.A. where it is a speciality in itself.

Urology deals with kidney disorders, including malignancies, bladder and ureter problems, kidney stones, male infertility and reproductive disorders, and diseases and malignancies of the prostate gland in males.

Transplant Surgery is a sub-speciality of General Surgery (except of the heart, which part of cardiac surgery) and involves the transplantation of live organs from donors e.g. liver, kidney, pancreas etc.

Therefore it can be seen General Surgery and the Surgical Subspecialties require a range of technical skills from microsurgery (e.g. small blood vessel anastomosis) to macrosurgery (e.g. bone cutting). It is important to develop assessment tools for surgical technical skills which can assess generic and specific technical skills, the range of technical skills required by a competent surgeon.

1.14 Surgical Pathway

Surgical operations in the U.K. are classified as elective, scheduled, urgent and emergency according to the NCEPOD report 1988 (National Confidential Enquiry into Perioperative Death). This classification is widely used. However in 2004 NCEPOD revised this classification into immediate (life threatening or other), urgent, expedited and elective. This reclassification is not yet widely used.

Intra-operative care involves several members of the surgical team. The surgeon determines the timing of the operation, the techniques, and the instruments and supplies to be used. The anaesthetist controls the patient's pain and, if necessary, the level of

unconsciousness to make surgery more tolerable and ensure that the patient regains consciousness safely and quickly following the operation. The scrub nurse readies all instruments, ensures the sterility of the surgical field, and anticipates when instruments will be needed by the surgeon. The circulating nurse makes sure the operating room is adequately supplied and provides any additional supplies to the scrub nurse during the operation. Postoperative care begins in a recovery room and continues in the ward, high dependency unit (HDU) or the intensive care unit (ICU) if the patient is very ill or needs intensive post-operative care with ventilation.

1.15 Evolution in Surgery

New techniques continue to advance the field of surgery. Lithotripsy, uses high frequency sound waves which are directed at kidney stones to break them so that they can be eliminated through the excretory system. Cryosurgery freezes and destroys abnormal tissue such as some cervical disorders, and to remove certain skin growths. Laser surgery uses a beam of light to vaporize or destroy tissue, a procedure commonly employed in ophthalmology, otolaryngology and neurosurgery.

Transplantation surgery, in which organs or tissues are removed from one person and surgically implanted into another person, is now performed for the eye lens and cornea, blood, bones and bone marrow, heart, lung, liver, and pancreas. In severe burn cases, healthy skin from an uninjured site is transferred to an area that has been damaged.

Recently the use of robots controlled by surgeons in performing heart surgery has enabled such operations to be done without a large chest wound, which keeps the patient post-operatively in hospital for a few days. Finally with smaller robots being made each day, in the future miniature or microscopic robots may be inserted in patients and microsurgery could be performed in arteries, in abdomen, in the chest etc.

Surgery is constantly evolving with new inventions in surgical techniques and methods. So surgeons must constantly learn new surgical technical skills. It is therefore important that we keep up with advancing surgical science and be able to assess each new surgical technique and technical skills required in a methodical and structured way.

1.2 Surgical Outcome

1.16 Overview

Surgery brings enormous benefits to patients and the outcomes of many common surgical operations are predictable, and for the most part, very good. However, complication rates may still be high, even for common operations and both mortality and morbidity can be very high for some major complex operations. In this chapter we review outcome data for common routine general surgical operations, consider the variability between units and between surgeons and the potential reasons for this. We also consider surgery in the broader contexts of studies of adverse events in healthcare and the understanding that the patient safety perspective has brought to this area of work. The chapter outlines that while considerable attention has been paid to the role of patient factors and co-morbidities in the prediction of outcome, comparatively little attention has been paid to the technical skills of the surgeon and other factors that may determine patient outcome.

1.17 Outcome of Common General Surgical Operations

In this section we will consider various measures of surgical outcome in the three common surgical operations we have chosen to evaluate technical skills in the thesis: open primary inguinal hernia repair, primary sapheno-femoral junction ligation and laparoscopic cholecystectomy. We will only concentrate on meaningful and salient measures of outcome in these operations. Other measures of outcome are summarised in the Appendix Table 1.1.

Open Primary Inguinal Hernia Repair

We will discuss the outcome of the open technique of primary inguinal hernia repair as this is the commonest technique used. The main measures of outcome for this operation are: operating time, recovery time, post-operative groin pain and recurrence rates of the hernia.

Operating Time

This is the time to complete the operation by the surgeon. Operating time is the easiest parameter to measure and can be used as a marker of technical expertise (Chapter 2). It has been used for the assessment of technical skills in open and laparoscopic surgery, on bench models and virtual reality simulators (Rosser et al 1997, Hanna ² et al 1997, Joice ¹ et al 1998, Starkes et al 1998, Feldman et al 2004). Liem M.S.L. et al 1997 reported in a series of 507 open inguinal hernia repairs with a mean operating time of 40 minutes (range 30-45 minutes). The operations had been performed by 87 Consultant and Trainee Surgeons, who all had significant experience in performing the operation.

Recovery Time

This is the total time for the patient is able to recover from the operation and is divided into hospital stay days, time to resume normal activities, time to return to work and time to resume athletic activities.

Putnis S. et al 2004 reported that 93.5% of all patients went home on the same day after hernia surgery. In one study normal activities were achieved in 5 days (Neumayer L. et al 2004) but in another it was achieved in 10 days (Liem M.S.L. et al 1997). The later study reported that patients returned to work after 21 days and to athletic activities after 36 days. Technical skills can affect recovery time. For example if there is post-operative bruising of

the tissues (the surgeon was not dextrous in handling the tissues) around the inguinal hernia repair site, recovery is slower for the patient.

Pain

This is pain experienced over the groin area where the inguinal hernia was repaired. This is divided into post-operative pain within 24 hours, after 7 days and after 30 days (chronic pain). The first two types of pain are related to the incision in the groin. Liem MSL et al 1997 reported using a visual-analogue score that the mean pain score for patients was 45/100 after 24 hours, 0 denoting no pain and 100 unbearable pain, in their large series of inguinal hernia repairs.

Chronic pain which may be due to neuropraxia of the ilio-inguinal nerve occurring at the time of the operation, and occurs for more than 30 days post-operatively and can last up to one year. Liem MSL et al 1997 reported a 14% rate and Neumayer L. et al 2004 have recently reported a 14.3% rate of chronic groin pain of surgery.

Recurrence

This is the recurrence of the hernia after the operation has been done to repair the hernia. For open inguinal hernia repairs the recurrence rate is less than 5 percent according two large studies in the late 1990's (Amid PK et al 1996, Kark AE et al 1998). This was recently confirmed to be a rate of 4.9% after two years after the operation by Neumayer L. et al 2004.

Primary Sapheno-Femoral Junction Ligation

The main measures of outcome for this operation are: operating time, deep vein thrombosis, quality of life years and recurrence rates.

Operating Time

This is the time to complete the operation as previously described. Chetter IC et al 2006 have reported a mean operating time of 48 minutes (range 41-63 minutes) for this operation.

Deep Vein Thrombosis

This is a potentially fatal complication of this type of operation. A clot forms in the deep veins and if not treated could dislodge and propagate to form a pulmonary embolus which can be fatal. Van Rij ² AM et al 2004 have reported a 5.3% rate of deep vein thrombosis after this type of surgery, and half of these patients being asymptomatic.

Quality-Adjusted Life Years

This means an individual's overall sense of well-being. In medicine, quality-adjusted life years is measured using various standardized factors as pain, treatment side-effects, mood, energy level, family and social interactions, sexual function, ability to work, and ability to keep up with routine daily activities. Quality of life is a combination of psychological and physical factors affecting the patient. Michaels JA et al 2006 have reported a significant improvement of quality of life, 2 years after this type of surgery.

Recurrence

Van Rij ¹ AM et al 2003 have reported a 32% recurrence rate of varicose veins after a technically successful first operation. Winterborn RJ et al 2004 however have reported a reduction of 60% recurrence rate after 11 years if the long saphenous vein is stripped instead of being left in-situ, which is sometimes done.

Laparoscopic Cholecystectomy

The main measures of outcome for this operation are: operating time, conversion to open operation, bile duct injury and recovery time.

Operating Time

This is the time to complete the operation as previously described. Jain PK et al 2005 have recently reported a mean operating of 41 minutes and 47 minutes for Consultant and Trainee surgeons respectively in a series of 269 patients having this type of operation.

Conversion to Open

This is normally when for technical reasons the operation can not be achieved laparoscopically, and a conversion to the open technique is required. Jain PK et al 2005 have reported an incidence of 1%, however in a large series of 6985 cases the conversion rate was 5.1% (Tarcoveanu E et al 2005).

Bile Duct Injury

This iatrogenic injury is considered to be a significant catastrophe in performing laparoscopic cholecystectomies. Its incidence has increased in the laparoscopic era compared to the previous open technique (Connor S & Garden OJ 2006). It is associated with significant morbidity and mortality (Moossa AR et al 1990, Savader SJ et al 1997). The incidence of bile duct injuries range from 0.1 to 0.5 per cent (Fletcher DR et al 1999, Cohen MM et al 1996, Russell JC et al 1996, Richardson MC et al 1996, Flum DR et al 2003).

Recovery Time

This is the total time for the patient is able to recover from the operation as previously described. McMahon AJ et al 1994 in their study of 302 randomised patients showed that laparoscopic cholecystectomy had a better recovery period than the open technique. Recovery after surgery was assessed by length of hospital stay, outpatient review at 10 days and 4 weeks, and patient questionnaires 1, 4, and 12 weeks after surgery. Median post-operative hospital stay was 2 days shorter after laparoscopic cholecystectomy. Laparoscopic patients returned to work in the home sooner; at 1 week, they had better physical and social functioning, were less limited by physical problems, and had less pain and depression. Jain PK et al 2005 have recently reported that 95% of patients in their study went home on the same day of surgery.

Technical skills can affect patient recovery time. For example if the ports are inserted not dextrously by the surgeon, the abdominal wall where the port is sited can be bruised and even a haematoma can form. As a result the patient recovers slowly.

1.18 Adverse Events in Surgery

As we can see from the previous section even in routine surgical operations adverse events can occur. This is clearly demonstrated in the increase of iatrogenic bile duct injuries in the laparoscopic era in performing laparoscopic cholecystectomies compared to the previous standard open technique.

Iatrogenic effects of treatments have been recorded in many studies, but only recently has the scale of injury to patients become apparent. The Harvard study (Brennan et al 1991) found that patients were unintentionally harmed by treatment in almost 4% of admissions in New York State. For 70% of patients the resulting disability was slight or temporary, but in

7% it was permanent and 14% of patients died partly as a result of their treatment (Brennan et al 1991). Serious harm therefore came to about 1% of patients admitted to hospital. Similar findings were reported from the Colorado and Utah study (Thomas ² & Brennan 2001). A parallel Australian study (Wilson et al 1995) found a 16.6% adverse event rate, with about half the cases being judged preventable, though with a similar number of serious incidents to the United States studies. A United Kingdom study (Vincent ¹ et al 2001) carried out a review of 1014 records finding a 10.8% adverse events rate, again about half were preventable. Emerging findings in New Zealand and Denmark also suggest a relatively high rate of adverse events around 10%.

A significant percentage of these adverse events are associated with a surgical procedure. The Harvard study found 47.7% and the Australian study found 50.4% of adverse events were associated with a surgical operation. The Utah Colorado Medical Practice study provided additional data on operative events (Thomas ² & Brennan 2001). The annual incidence rate of adverse events among hospitalised patients who received an operation was 3.0%. Among all surgical events, 54% were preventable. Eight operations were identified as high risk, based upon their preventable adverse event rate: lower limb vascular bypass graft (11.0%), abdominal aortic aneurysm repair (8.1%), colonic resection (5.9%), coronary artery bypass graft or cardiac valve surgery (4.7%), transurethral resection of the prostate or of a bladder tumour (3.9%), cholecystectomy (3.0%), hysterectomy (2.8%), and appendectomy (1.5%). Technique-related complications, wound infections, and postoperative bleeding produced nearly half of all surgical adverse events.

A recent Australian study (Kable et al 2002) found the rate for adverse events in surgical admissions was 21.9%. Disability that was resolved within 12 months occurred in 83%, 13% had permanent disability, and 4% resulted in death. Reviewers found that 48% of

adverse event were highly preventable. The risk of an adverse event depended on the procedure and increased with age and length of stay.

In the UK complication rates for some of the major operations are 20-25% with an acceptable mortality of 5-10% (Gordon et al 1993). However at least 30-50% of major complications occurring in patients undergoing general surgical procedures are thought to be avoidable (Healey et al 2002).

The wide variation in surgical complication rates between different centres and different surgeons would support this view. Many adverse events classified as operative are, on closer examination, found to be due to problems in ward management rather than intra-operative care. For instance Neale et al 2001 identified preventable pressure sores, chest infections, falls, poor care of urethral catheters in their study of adverse events, together with a variety of problems with the administration of drugs and intravenous fluids.

As we can see adverse events vary in their rates depending on different institutions. The preceding section also demonstrates the variability of outcome measures in routine general surgical operations e.g. recurrence rates, bile duct injuries etc. These adverse events can be related to technical skills errors by the surgeon. Therefore by assessing surgical technical skills one may predict or expect adverse events to occur or not. The next section will outline patient and non-patient factors which determine the outcome of a patient after surgery.

1.19 Determinants of Surgical Outcome

There are several factors affecting patient outcome. These include *patient factors* e.g. age, weight etc *non-patient factors* e.g. time of operation, surgeon and organisational factors

e.g. availability of HDU or ICU bed. These are summarised in the Appendix in Tables 1.2, 1.3, 1.4 and 1.5. All these factors interact with one another and affect the overall patient outcome.

In this section we will outline patient, surgical technical skills and other non-patient factors which all can contribute to the overall outcome of the patient after surgery.

Patient Factors

The patients' own factors and co-morbidities can determine the surgical outcome. These can be related to directly to the patient e.g. age, sex, BMI etc or disease processes affecting the patient e.g. cardio-respiratory disease, renal function, extent of the disease requiring surgery etc. A complete list of patient factors affecting outcome are summarised in Appendix Table 1.3.

Over the last few years there have been various scoring systems constructed in the aim to predict surgical patient outcome. Table 1.20 summarises some of the scoring systems which have been devised over the years.

Table 1.20 – Surgical Outcome Prognostic Scoring Systems

Surgical Outcome Prognostic Scoring Systems
APACHE II
POSSUM
Modified POSSUMS
E-PASS
Hardman Index

The APACHE II scoring system is a general measure of disease severity based on, current physiologic measurements, age, and previous health condition. Scores range from 0-71, an increasing score is associated with an increasing risk of hospital death. It is usually used in the intensive care unit setting. Appendix Table 1.6 summarises the factors which contribute to an APACHE II score. This scoring system has been shown to predict patient outcome (Knaus & Draper 1995). However it is used more by intensive care doctors than surgeons.

The original POSSUM (Physiological & Operative Severity Score for the enUmeration of Mortality and Morbidity) score was devised by Copeland ¹ et al 1991 in Liverpool, UK.

It measures surgical outcome depending on several factors including: physiological status of the patient, disease process that requires surgical intervention, nature of operation, and pre and peri-operative support. Raw morbidity and mortality data alone can provide a biased picture in the relation to patient outcome. POSSUM however, allows risk-adjusted assessment of surgical quality. It can accurately predict 30-day morbidity and mortality. The factors are summarised in Appendix Table 1.7. POSSUM has been modified to P-POSSUM (Prytherch et al 1998) and for individual surgical sub-specialties (Law et al 2005, Tekkis et al 2004). There are other prognostic outcome scoring systems e.g. E-PASS (Haga et al 2004) and Hardman index (Neary ² et al 2003), but these are not widely used in the surgical community to predict patient outcome.

Surgical Technical Skills

The technical skills of the surgeon are fundamentally important to the outcome of a patient after surgery. However despite this, technical skills are very rarely assessed during actual operations whether for training or research purposes (Darzi A. et al 2001). The primacy of technical skills is nevertheless the underlying assumption and is implicit in the creation of league tables or rankings of surgical performance (www.drfooster.co.uk). Once outcomes

(usually mortality) have been correctly adjusted for patient risk factors, the remaining variance is presumed to be explained by individual surgical skill.

In the Bristol Case, technical skills in cardiac surgery were closely linked to surgical outcome (www.bristol-inquiry.org.uk). The final report in 2001 concluded that the mortality rate in the Bristol Paediatric Cardiac Unit was higher than the national level and was due in part to poor technical skills of the cardiac surgeons involved. The surgeons themselves did not pick this up and was only investigated after another healthcare professional was concerned on the higher than normal mortality rates. This was a high profile case and was a major watershed in surgical practice in the UK.

There have been previous concerns on a high level of mortality in a short period of time, De Leval et al 1994 reported a cluster of deaths following paediatric arterial switch operations which was due to surgical performance. The cardiac surgeon who reported the cluster, re-trained and as a result the mortality rates decreased. In this case the surgeon involved was able to reflect and appraise his own technical skills, unlike the Bristol surgeons. However not all surgeons declare or investigate if they have a high mortality or morbidity rate. In 1998 an anonymous article by a junior doctor speculated (Anonymous et al 1998), a now retired surgeon had a higher rate of mortality than normal in his or her opinion. This junior doctor reported this to another Consultant Surgeon, but nothing was done about this. If surgical outcome is to be monitored it should be done in an unbiased and systematic way.

There are several factors affecting the technical skills of a surgeon. These can be related to patient factors, some of which are summarised in Table 1.2 & 1.3. For example obese patients are more difficult to operate on, as the dissection planes are not as well defined as

thin patients. Also patient's pathology can affect technical skills e.g. a fixed colonic tumour is technically harder to operate on than a mobile one, again due to undefined dissection planes.

Surgeon's factors can also affect the technical aspects of the operation or procedure. These are: Experience, Ability and Human Factors. I will now give a brief outline of these factors which affect the technical skills of a surgeon and therefore patient outcome.

Experience

Inexperience or, lack of training also contributes to poor surgical outcome. Porter GA et al 1998 looked at variations in outcome between patients with rectal cancer treated by specialist vs. non-specialist colorectal surgeons, and independent of that, between results of surgeons with high vs. low volume work. The analysis showed that the risk of local recurrence was increased, and disease specific survival was lower, in patients treated both by surgeons not trained in colorectal surgery and by surgeons performing less than 21 procedures during the study. Thus the best result was obtained from the trained surgeon performing higher volume work (10.4% recurrence, 67.3% survival) and the worst by the untrained surgeon with low numbers (44.6% recurrence, 39.3% survival). Birkmeyer et al 2002 analysed 14 types of procedures and demonstrated mortality decreased as volume increased with Begg et al 2002 demonstrating similar findings for morbidity rates in radical prostatectomy.

Ability

The technical ability of the surgeon is also important to the surgical outcome of the patient. For example, normally a perforated duodenal ulcer is repaired by an open operation. Due to the incision being higher up in the abdomen, the patient tends to suffer more pain and is in

hospital for 5-7 days. However if the surgeon has the ability to do a laparoscopic repair, the patient tends to have less pain and go home after 1-3 days later. This has been demonstrated in studies comparing open and laparoscopic repair of perforated duodenal ulcers (Lunevicius R & Morkevicius M. 2005, Tsumura H et al 2004). This better recovery period has also been demonstrated in advanced laparoscopic colorectal surgery (Kariv Y et al 2006, Alves A et al 2005). A surgeon's ability increases with more operative surgical experience. He or she accumulates generic and specific technical skills through learning different technical skills. For example a junior surgeon learns how to do an open appendicectomy first, but after acquiring laparoscopic technical skills he may do the appendicectomy laparoscopically. However surgeons are heterogeneous and some surgeons have more technical skills ability than others (Wanzel² KR et al 2002).

Human Factors

These are factors which can affect human performance (Bogner MS 1994). Surgeons performing a task or operation are subject to these factors. They are:

Physiological Factors:	Fatigue, sleep, hunger, state of health
Emotional Factors:	Mood states (e.g. anxiety), life events, confidence, motivation, personality
Cognitive Factors:	Thought processes involved in tasks such as planning and decision-making (cognitive biases)
Stress Factors:	Difficult or novel tasks, work load, interruptions, urgent tasks, etc

It has been recently shown that these stress factors can affect surgical performance, and may be detrimental to outcome in a questionnaire based study (Wetzel CM et al 2006). However there is no current study evaluating the effects of stress on technical skills on real live operations and the subsequent surgical performance.

This current section has outlined how technical skills and the factors influencing it can affect surgical outcome. In the following section other non-patient factors not related to the surgeon directly, which can influence patient outcome, will be outlined.

Organizational & System Factors

An operating theatre is an extraordinarily complex system. The complexity is manifested not only in the patient and their condition, but also in the sophistication of instrumentation, the high volume of information that must be processed, the nature of communication and team co-ordination, and the urgency and occasional uncertainty with which decisions and interventions must be made. This complexity combined with heavy workloads, fatigue and production pressures, makes surgical care particularly vulnerable to adverse events, therefore poor outcome. Despite this vulnerability most cases are performed proficiently and safely, highlighting the resilience of individuals and surgical teams to the potential adversity of the setting. In the following section we will outline non-patient factors which can affect patient outcome.

Calland et al 2002 have argued for a 'systems approach to surgical safety'. Such approaches suggest that it is necessary to study all aspects of the system that comprises a surgical operation, including such issues as equipment design and use, communication, team co-ordination, factors affecting individual performance and the working environment.

Surgical adverse events (poor outcome) may be due to poor communication, bad operative technique, malfunctioning or improperly used equipment, cognitive errors due to stress or inattention, all compounded by resource and organisational problems. However these factors have been poorly studied in the field of surgery. Communication in the operating

suite is often poor and may be implicated as a contributor to adverse outcomes (Helmreich RL & Schaefer H. 1994, Runciman WB et al 1993, Sexton JB et al 2000).

The outcome of care is, broadly speaking, determined by the structural aspects of the system (personnel, environment, infrastructure) and the process (the procedures carried out the quality of care provided) Donabedian A 1968. The major contributory factors to the quality and safety of care are outlined in Appendix Table 1.8. The higher level influences (institutional context and organisation and management) cannot be directly assessed, but the remaining five levels are all of considerable importance: patient factors, task factors, individual factors, team functioning and working environment.

The assessment of team performance in surgery is embryonic and a major task for the future is to develop team performance measures which are both grounded in more general understanding of team dynamics, yet cognisant of the particular features of the operative environment.

In the present state of knowledge, we do not know which factors are most critical to outcome, beyond patient risk factors, basic levels of individual skill and the organisation of care. The factors that lead to very high quality performance may differ from those that are required to achieve an average level of performance. For instance surgeons, who routinely achieve good outcomes in even adverse circumstances, may be distinguished more by their pre-operative and intra-operative clinical decision making than their technical skills.

1.20 Conclusions

As one can see good surgical patient outcome is fundamental to good surgical practice. There are many factors affecting patient outcome after an operation or procedure, which can be patient or non-patient related. The important outcome measures vary in individual operations, and even these vary depending on the surgeon & institution. Adverse events after surgery has been shown to be detrimental to the patient and incidence of them also vary depending on country, surgeon and institution. Finally non-patient factors have been shown to influence the outcome of patients after surgery. These can be directly related to the surgeon (technical skills) or not e.g. team performance.

In this thesis we will examine only one non-patient factor which can affect patient outcome: the surgical technical skills of the operating surgeon. Although we will not do any outcome studies in the thesis we aim to demonstrate that technical skills is important in the quality of the operation being performed and is therefore inextricably linked to surgical patient outcome. This is because one must first construct reliable and valid assessment tools in surgical technical skills before this aspect can be correlated to patient outcome, with meaningful results. After constructing such tools in technical skills assessment we aim to eventually correlate this to patient outcome. However this is beyond the scope of this present thesis, and will be conducted in the future as a study.

1.3 Training in Surgery

1.21 Overview

Surgical training has changed in the U.K. in the last decade. There are now yearly appraisals called record of in training assessment (RITA). Although this structured approach to appraisal is welcome, there is still no structured assessment of technical skills and ability of the trainee.

This is assessed by unstructured direct observations by the surgical trainer and an up to date operative logbook. These methods however have shown to be an unreliable way to assess surgical technical skills and competency. The acquisition of technical skills remains to be done in an apprenticeship style of learning. However this method of learning is unlikely to be effective in the long term as the surgical trainees working environment has changed recently. This chapter outlines the history, present day, different modes and the future of surgical training.

1.22 History

The Roman Galen was a strong advocate of cadaveric dissection in the acquisition of surgical skills. Hindu surgery was well advanced by the time of Susrata in the 5th century AD, who also advocated dissection of dead bodies for the practice of technical skills. On the ensuing centuries surgery did not attract the intellectual kudos as physicians. As a result by medieval times surgery was being performed by barbers. Surgeons tended to be slightly more intellectual than their barber counterparts, but mainly came from the lower classes. They learnt their craft by passing the skills from father to son.

In 1540 Henry VIII tried to improve the standards of surgery by forming the Barber-Surgeon company as a result of an ongoing dispute between the two factions. Strict rules

were laid down, and an examination was introduced at the end of each surgeon's apprenticeship. This was the first form of formal assessment in surgical training. In 1745 the company was dissolved and up to the 19th century surgical training was viewed as a simple apprenticeship.

John Hunter promoted surgery as a scientific field, by teaching his students anatomy, physiology and pathology as part of their surgical training. As a result many of his students practised and disseminated his theory of surgical training to other parts of the U.K. and the United States. William Halstead also greatly contributed to the training of surgery at around the same time. He emphasised that tissues should be handled gently with minimal blood loss, and the close relationship of physiology and surgery. He set up a surgical residency programme based on German theories of surgical training. This idea was promoted by himself and his students across the United States. As a result by 1928 the American Medical Association approved the principles of the residency programme and its principles still hold to this day. In the U.K. and the rest of Europe surgical training was done by an apprenticeship system and reputations were promoted by publications. Throughout the 20th century there was a lack of a methodical and systematic approach to surgical training in the U.K. In 1972 at Guy's Hospital, London there was an attempt to train surgical registrars based on adequate range and volume of surgery. It was appreciated at that time that assisting in operations did not confer the type of acquisition of knowledge and skill to perform an operation.

1.23 Present Day

Prior to Calminisation in the mid 90's surgical training could take a long time to achieve competency in the apprenticeship style of training. During the registrar years of a surgical trainee he or she took the Fellow of the Royal College of Surgeons exam from any of the

four surgical colleges (England, Edinburgh, Glasgow or Ireland). It consisted of two parts. The first part was the basic sciences of anatomy, physiology and pathology. The second part was on clinical surgery and the knowledge of operations. There was no formal assessment of surgical technical skills.

Calman merged the registrar and senior registrar positions into the Specialist Registrar grade. Basic surgical training was done at the Senior House Officer (SHO) level with a compulsory basic surgical skills course and assessment. With the advent of the new MRCS exam, such aspects as communication skills were assessed. After passing this exam one could apply for a Specialist Registrar position and enter into higher surgical training in the trainee's speciality.

Higher surgical training consists of yearly record of in training assessment (RITA), with successful completion of each year a prerequisite for satisfactory progression. The appraisal system does not formally assess surgical technical skills. Although the basis of this was to try to assess surgical trainees in a structured manner it is widely considered to be inadequate.

1.24 Modes of Training

Ideally operating on real live patients would be the optimal mode for assessment for surgical training. However there are ethical and clinical safety issues surrounding surgeons, especially junior surgeons to learn basic surgical techniques on live patients. However for senior surgical trainees need to operate and be assessed on live patients. Assessing technical surgical skills in the laboratory setting is a safe and controlled environment for surgical trainees, especially junior surgeons to train and possibly be assessed on. In the

laboratory setting each technical task can be reproduced and immediately repeated, broken down into several components, and have instant self and mentor feedback.

Inorganic – Synthetic

Synthetic models are divided into non-life and life-like. They are particularly useful in teaching and learning basic and simple surgical technical skills. Such models are used in the compulsory Basic Surgical Skills course as well as other courses. Simple jigs can be used for knot tying and suturing techniques, and if enclosed in a laparoscopic box trainer, laparoscopic techniques can be learnt and practised e.g. laparoscopic suturing. Life-like models use various latex and rubber materials to construct a model. They can be high or low, fidelity reflecting the realness of the model to the reality. Low fidelity models are cheaper and are useful in practising simple surgical techniques. High fidelity models are more expensive and some models replicate a complete operation e.g. sapheno-femoral junction ligation with a local circulation connected to it. Other operations like aneurysm repair, femoro-popliteal bypass, fracture fixation, joint replacement, pelvic surgery etc can also be reproduced.

Inorganic – Electronic

Computer or Virtual Reality Simulation

There is a marked increase of this type of simulation for surgical training. With increasing higher fidelity and haptics (tactile feedback) some simulators are becoming very realistic. The main advantages of them are: a safe environment, real time performance, record and playback, repetition and practice, immersive environment, objective measures of performance, impartiality, instant feedback and results. Also with costs coming down and computers being more powerful, simulators are becoming more affordable for institutions to buy. Simulators in use for example are for endoscopy (upper gastro-intestinal, colonoscopy, bronchoscopy etc) and laparoscopy (arthroscopy, cholecystectomy, suturing

etc). Any system should undergo extensive validation (Reznick ² 1999) before they are used in training and assessment, as they can be shown to be inappropriate (Smith S. ¹ et al 1999).

Organic - Animal

Current U.K. law does not permit the use of live animal tissue to be used to practice surgical techniques, unlike in Europe, U.S.A. and other countries. Worldwide a number of live animal models have been developed. Live animal models are expensive and their anatomy can greatly vary from humans. However despite this they are used widely and provide live simulations that mimic the real operative reality. Examples include the dog in the coronary bypass model, also training vascular trainees in bypass procedures in dogs have shown to improve technical complications and patency rates (Saifi et al 1990). Other models include laparoscopic cholecystectomy in pigs but the pig's anatomy does vary from humans at Calot's triangle. The use of dead animal tissue is much cheaper and does have less ethical issues. Also it is permitted in the U.K. to be used. Sheep and pig material have been used for bowel anastomosis in surgical skills courses for example.

Organic - Cadaver

The main advantage of this type of training is that it has the correct anatomy and relations. Cadavers are widely used in the U.S.A. and some courses in the U.K. are entirely cadaveric dissection e.g. Definitive Surgical Trauma Skills® course. There is however always a short supply of cadavers, and there are also ethical issues surrounding such dissection. Also formalin preserved human tissue behaves differently from live human tissue and cannot reproduce bleeding.

Organic - Live

Obviously operating on live patients is the optimal way of training. However with increased public scrutiny and implications of clinical safety it is probably not appropriate

for very junior surgeons to train initially on live patients and time spent perfecting surgical skills in the laboratory setting is probably more beneficial. After acquiring basic surgical skills junior trainees under supervision should then operate on live patients. Senior trainees must learn and train on live patients, but also time spent in the laboratory on animal, cadaveric dissection etc complementing live operating again would be beneficial for the trainee.

Surgical Courses

The advantages of surgical skills courses include close supervision, self-instruction, feedback and the repeated practice of techniques and procedures. A skills laboratory is useful in introducing the complex aspects of fine motor skills in a controlled environment (Barnes¹ 1987 & Barnes² et al 1989, Lossing et al 1992). Surgical skills courses must complement normal hands on training in the normal clinical environment (Cuschieri¹ 1993, Kirk 1996).

Basic Surgical Training

Since 1996 in the UK all basic surgical trainees must attend and successfully complete a Basic Surgical Skills course before they are allowed to sit the MRCS exam. The course is run by the Royal College of Surgeons with courses being held widely in hospitals ranging from teaching to district. The course discusses and teaches simple basic surgical skills like knot tying, types of knots, laparoscopic skills etc. Synthetic and animal tissues are used to simulate real human tissue.

Higher Surgical Training

The Royal Colleges and individual institutions have developed various courses for the surgical sub-specialties, these are not compulsory and use a combination of synthetic,

animal and cadaveric tissues. In general surgery for instance there is a specialist registrar course in general surgery aimed at year 1-3 higher surgical trainees, teaching safe surgical techniques e.g. types of anastomoses, stoma formation, laparoscopic surgery skills.

Then there are sub-specialities courses which are aimed primarily at more complex training. There are colorectal, breast, vascular and laparoscopic courses specialising in specific operations e.g. anterior resection, breast reconstruction, femoral bypass, laparoscopic cholecystectomy.

1.25 The Future

The future of surgical training is at present changing. With new technologies and new research in this field we may be able to use these to implement a structured method of teaching and learning technical skills in surgery. Also if we can integrate all these methods of teaching and learning into a curriculum of acquiring technical skills, this would be beneficial to surgical trainees, as well as the future of surgery. This thesis aims to develop tools which can be used to assess technical skills and can possibly be used in a structured curriculum for surgical training.

Chapter 2

Assessment of Technical Skills in Surgery

2.10 Overview

To acquire surgical competency involves the surgeon to be proficient in a combination of technical skills, decision making (pre-operative, operative, post-operative), team performance which are all underpinned by knowledge and communication. Performing complex technical skills procedures daily differentiates surgeons from most physicians in their normal daily working practice. Assessing the technical skills in surgery is challenging and complex. This chapter provides a review of the assessment methods used to assess technical skills in the last 30 years to date.

2.11 Background

Whilst surgical knowledge, clinical skills, communication are assessed by examinations in the UK such as the MRCS (parts 1 & 2) and the FRCS (part 3) conducted by the Royal College of Surgeons in the U.K., there have so far been few attempts at assessing surgical technical skills on a formal basis, due to the lack of structured methods of assessments.

The typical surgical learning is based on an apprenticeship model under the supervision of a surgical trainer (Consultant Surgeon). In this model the assessment of technical skills competency is the responsibility of the trainer, and a logbook is used. However it has been shown that this method can suffer from poor reliability of assessing performance (Chatenay et al 1996). To ensure reliability, assessment techniques should be as objective as possible. Table 2.1 summarises the qualities of a good assessment tool (Gallagher ² et al 2003). All of these apply if one wants to develop an assessment tool in technical skills in surgery, so it is widely accepted and used.

Table 2.1 – Principles of an Assessment Tool

Reliability is a generic term that covers all aspects of the dependability of a measurement test or tool. It is the internal consistency, or the extent to which the assessment tool yields the same results repeatedly under similar conditions. For likert and checklist scoring, intra-class correlation coefficient and kappa coefficient are important. Reliability is an important aspect for this thesis.
Validity is the fundamental aspect of any measuring/assessment instrument/tool and that it measures or assesses what it claims to measure. Within the testing literature, a number of validation benchmarks have been developed to assess the validity of a test or assessment tool. These include face validity, content validity, construct validity, concurrent validity, discriminate validity, and predictive validity. All of these validities are important for this thesis.
Face Validity is defined as a type of validity that is assessed by having experts review the contents of a test to see if it seems appropriate for the test or assessment.
Content Validity is defined as an estimate of the validity of a testing or assessment instrument based on a detailed examination of the contents of the test or assessment items. Establishing content validity is also a largely subjective operation and relies on the judgements of experts about the relevance of the materials used.
Construct Validity is defined as a set of procedures for evaluating a testing or assessment instrument based on the degree to which the test items identify the quality, ability it was designed to measure. A common example is the ability of an assessment tool to differentiate between experts and novices performing a given task or operation.
Concurrent Validity is defined as an evaluation in which the relationship between the test or assessment scores and the scores on another instrument claiming to measure the same construct are related.
Discriminate Validity is defined as an evaluation that reflects the extent to which the scores generated by the assessment tool actually correlate with factors with which they should correlate. An example of this would be an assessment tool that could differentiate levels of ability within a group with similar experience, such as discriminating abilities of all the junior registrar trainees.
Predictive Validity is defined as the extent to which the scores on a test or assessment are predictive of actual performance. An assessment tool used to measure surgical technical skills will have predictive validity if it predicts who will perform actual surgical tasks or operations well and who will not.

Trainees who enter a surgical training programme (e.g. Basic Surgical and subsequently Higher Surgical Training) have already acquired a certain amount of skill by observing

other surgeons and have developed certain methods of performing tasks, sub-tasks and technical skills decision making. One important aspect of self-directed learning is the need for continuous evaluation and feedback. Feedback is essential because deficiencies in training and performance are difficult to correct without feedback.

The absence of structured formative feedback methods and the increasing attention of the public and media on the performance of surgeons have given rise to an interest in the development of structured methods of assessment of surgical technical skills. However these methods are still in their infancy. Table 2.2 summaries the current methods of assessing surgical technical skills. Table 2.3 summarise the new methods being used to assess technical skills in surgery.

Table 2.2 – Current Methods of Assessment of Technical Skills

Method of assessment	Reliability	Validity
Procedure lists with logs	Not applicable	Poor
Direct observation	Poor	Modest
Direct observation with criteria	High	High
Animal models with criteria	High	Proportional to realism
Videotapes	High	Proportional to realism

Table 2.3 – New Methods of Assessment of Technical Skills

Checklists
Likert rating scales e.g. OSATS
Dexterity analysis systems e.g. ICSAD, ADEPT
Virtual Reality Simulators
Analysis of the Final Product
Error Scoring

2.12 Method of Review

To review this subject in assessing technical skills in surgery a literature search was performed. Using the key words ‘assessment’, ‘competence’, ‘skills’, ‘surgery’, ‘technical’ and ‘training’ a PubMed/Medline search was conducted. Only original papers between the 1st January 1975 to 31st December 2005 on assessment of technical skills in surgery (open, endoscopic, laparoscopic) were included. Endoscopy and laparoscopic surgery are combined into a single category of MIS (Minimally Invasive Surgery) for simplification. Reviews, articles, editorials, commentaries and abstract publications and presentations were excluded from the review process.

Each paper was then analysed and summarised under the following categories: year of paper, main author, originating institution, number of subjects, type of subjects, methods and materials, results and conclusions from the study. The methods and materials section was then subdivided into the following sub-headings: type of assessment e.g. bench model, real patients, virtual reality simulator etc, method of assessment e.g. likert scoring, checklist, motion analysis etc, type of procedure e.g. open or minimal invasive surgery (MIS). Each paper was then assimilated into a chronological table.

Table 2.4 summarises chronologically all the 130 assessment of technical skills in surgery papers between 1975 to 2005. These studies on open and MIS surgery, are on bench models (75) (synthetic, animal and human cadaveric material), virtual reality simulators (35) and on real live animals (8) or humans (14). 89 studies had 20 or more subjects in their study. In each category we will sub-analyse these 89 studies.

The following chapter outlines each method of assessment used in these papers as well as methods used. Each method is divided either into generic or (procedure) specific technical

skills assessment. The following paragraph describes and explains generic and specific technical skills.

Generic and (procedure) specific technical skills are different skills which are acquired at different times during surgical training. By the time the trainee reaches Consultant Surgeon status he or she should have mastered these aspects of technical skills in surgery. Generic technical skills are mainly acquired during basic surgical training e.g. knot tying, dissection, handling of tissues, instrument handling. However more advanced generic technical skills are acquired during higher surgical training e.g. laparoscopic knot tying, micro-vascular suturing. These basic skills are required by the surgeon to perform a complete procedure. They alone do not form a task or sub-task but allows the surgeon to perform so.

Specific technical skills are learnt throughout basic and higher surgical training in an incremental fashion (easy-level 1, intermediate-level 2) e.g. open inguinal hernia repair, open hemicolectomy etc and advancing to (complex-level 3) e.g. anterior resection, laparoscopic fundoplication etc. These specific technical skills are essential to the quality and outcome of the operation e.g. correct insertion of mesh in open inguinal hernia repair or viable bowel anastomosis in anterior resection. They are a combination of key tasks which are divided into sub-tasks.

The following account of the methods of assessing technical skills are categorised into either non-structured or structured methods and then sub-divided to either assessing generic or specific technical skills.

2.13 Current Methods of Assessment of Technical Skills

Didactic Knowledge – Generic & Specific Technical Skills

The focus of surgical examinations such as the MRCS and FRCS is mainly on the communication and knowledge of the trainee, and only the theoretical knowledge of generic and specific skills of operations and procedures. However such examinations fail to assess the technical ability of a trainee. One study showed that there was no relationship between the American Board of Surgery in Training Exam (ABSITE) and the clinical skills in an OSCE examination (Schwartz et al 1995). The ABSITE score also does not correlate with technical skills (Scott¹ et al 2000). Thus examinations have poor content validity for the assessment of technical skills. Recently Schell et al 2004 confirmed that didactic teaching is not as good as hands on teaching in acquiring generic technical skills.

Time Taken for Surgery – Generic Technical Skills

Historically speed of performing operations was important, especially in pre-anaesthesia days. This is not as important now, as attention to detail in performing complex operations is more important, which could not have been done in the pre-anaesthesia era. However, as time is the easiest parameter to measure it is still used as a marker of technical expertise. It has been used for the assessment of technical skills in open and laparoscopic surgery, on bench models and virtual reality simulators (Rosser et al 1997, Hanna² et al 1997, Joice¹ et al 1998, Starkes et al 1998, Feldman et al 2004). It is also used for determining the learning curves for surgical and new procedures (Watkins et al 1996, Everbusch A & Grantcharov TP 2004). While operative speed is a desirable surgical quality, in order to lower the time spent under anaesthesia, it fails to assess the quality of surgical performance (Shah et al 2003). While it could be used as a measure of skill on standardised bench

models, it is an unreliable measure when used during real procedures, due to the influence of various other factors such as patient characteristics.

Operative Log Books – Specific Technical Skills

All surgical trainees are expected to maintain a log book of their operative experience which must be submitted at the time of surgical examinations (MRCS/FRCS), appraisal times and at the annual RITA (Record of in Training Assessment). Currently higher surgical trainees in General Surgery in the UK must submit computerised log books to the Association of Surgeons of Great Britain & Ireland annually so the number of procedures can be monitored. However, it has been found that they are merely indicative of procedural performance (Sandermann et al 2002) and not a reflection of operative ability and thus lack content validity. Galasko & MacKay's 1997 said log books are useful for unsupervised surgical training, but these comments are out of date in a Consultant Surgeon lead patient service in the UK. Current surgical trainees do not follow a set curriculum of acquiring technical skills and rely an ad hoc way of acquiring these (Sarker SK ¹ 2003).

Direct Observation by Surgical Trainers – Generic & Specific Technical Skills

The assessment of technical skills, as currently occurs in the operating room is subjective. As it is not based on specific criteria, the assessment is unreliable. It is not only influenced by the subjectivity of the observer and would have thus possess poor test-retest reliability, but also suffers from poor inter-rater reliability as even experienced senior surgeons would not possess the same level of agreement while rating the skills of a trainee. The Joint Committee for Higher Surgical Training in General Surgery in the UK has appreciated the need to increase the extent of emphasis laid on the assessment of technical skills in surgical training. With that intention it has recommended the use of the Operative Competence assessment with Consultant Surgeons assessing their trainees on procedures

performed during a fixed training period. The assessment consists of the trainee's ability to carry out procedures using 5 overall likert ratings (Table 2.5) (www.jchst.org). Although, this assessment method has face validity, there is only one study that has demonstrated its construct validity or reliability, presented as an abstract (Burt ¹ et al 2005) and not been published to date as a full paper. One study using this tool, found that the assessment of the Trainees by their Consultants was as accurate as the trainees self-assessment (Burt ² & Wilkins 2005).

Table 2.5 - Operative Competence- JCHST

U: Unknown (not assessed) during the training period
A: Competent to perform the procedure unsupervised (can deal with complications)
B: Does not usually require supervision but may need help occasionally
C: Able to perform the procedure under supervision
D: Unable to perform the entire procedure under supervision.

2.14 Structured Methods of Assessment of Technical Skills

Likert Scores and Checklists – Generic & Specific Technical Skills

Table 2.7 summarizes all 56 papers with 20 or more subjects in them using likert and checklist scoring. Both scoring systems are good in assessing generic technical skills, feedback and training improves the acquisition of these skills.

Checklists have been used across a number of specialities including open and MIS. The availability of set criteria against which technical skills can be assessed makes the assessment process more structured, valid and reliable. The Toronto group (Lossing et al 1992, Winckel et al 1994, Reznick ¹ et al 1997, Martin ² et al 1997, Anastakis ^{1,2} et al 1999 & 2003, Szalay et al 2000, MacRae et al 2000, Grober^{1,2,3} et al 2003 & 2004) have

extensively used checklists with likert scoring, and found that likert scoring has a higher inter-rater reliability as well as construct validity compared to checklists. The main problem with checklists are they are either finite 'yes' or 'no' answers, and when assessing the technical skills of Consultant Surgeons and Trainees this may not be sensitive enough. The likert scoring system allows more range in scoring when the particular technical skills question is being asked.

The success of the OSCE examination led to the development of a similar concept for the assessment of technical skills by a group in Toronto (Martin ² et al 1997). The OSATS (Objective Structured Assessment of Technical Skill) likert scoring assessment originally consisted of six stations where residents/trainees perform procedures on live animal or bench models in fixed time periods (Reznick ¹ et al 1997). The assessment of performance during the performance of tasks is done using an operation or task specific checklists and a likert rating scale (Table 2.6). The checklist consists of items scored as a yes/no, which are considered essential for the competent performance of the task at the relevant station. The likert scale consists of seven components of operative skill which are marked on a 5 point likert scale. A trainee's score is the combination of the score on the seven components. These seven components are generic to any operation.

Table 2.6 – Objective Structured Assessment of Technical Skills (OSATS)

	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 & 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 the names of most instruments and used appropriate instrument or the task		Obviously familiar with the instruments required and their names
Use of Assistants	Consistently placed assistants poorly or failed to use assistants		Good use of assistants most of the time		Strategically used assistant to the best advantage at all times
Flow of Operation & 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 aspects of the operation		Demonstrated familiarity with all aspects of the operation

Both live animal operating and bench models have been used for the OSATS assessment. There were no differences in the performance of trainees in both formats of the examination. Both were able to strongly discriminate between junior and senior residents, thus demonstrating the construct validity of OSATS (Martin ² et al 1997). The inter-station reliability for the six stations was 0.84 for the likert rating scale and 0.78 for the checklists (Reznick ¹ et al 1997). Regehr et al 1998 studied the correlation between the checklists and the likert rating scales. They found that the presence of checklists did not substantially improve the reliability or validity of the likert rating scale over that of the likert rating scale alone.

Using OSATS (likert scoring) as a measure of surgical skill, Anastakis ² et al 1999 found an effective transfer of skills from bench training models to a human cadaver. Datta ⁵ et al 2004 demonstrated a strong correlation between performance on bench models and during real operations with no difference in the likert rating scores for both platforms of assessment. The only drawback to the performance of the OSATS assessment are the

resources and time involved in getting a number of staff surgeons to observe the performance of trainees. These studies using OSATS are a good discriminator between inexperienced and experienced surgeons i.e. generic technical skills, but do not assess the quality of the operation i.e. procedure specific technical skills. The group has added 'quality of product' to the likert scoring system. However this as well as 'knowledge of procedure' does not comprehensively assess the specific skills needed to assess the final quality of the procedure or operation. Also most of these studies have been on bench models in laboratories or live animals. In the UK operating on live animals is not permitted.

It has been suggested that retrospective video watching of performance may be a way forward. It does not involve the presence of multiple faculty raters at the same time and also adds to the element of objectivity by making the assessment blinded (Beard ¹ et al 2005).

Table 2.7 – Likert & Checklist Scoring Papers (with 20 or More Subjects)

Number of Papers	Total Number of Subjects	Conclusions
56	2755	Likert & checklist have good reliability & validity. Likert scoring is a better discriminator and reliability than checklist scoring. Feedback & training improves technical skills

Dexterity Analysis Systems – Generic Technical Skills

Table 2.8 summarises the 28 papers with more than 20 subjects using dexterity analysis to assess technical skills. Motion analysis assesses generic technical skills i.e. time and motion. These skills tend to improve with increased experience and case number.

A generically technically skilled surgeon is one who executes a task quickly, is economical in his or her hand movements as well as being precise and accurate. While current methods

such as likert rating scales and checklists evaluate the technical ability of the trainee executing certain aspects of a task/procedure, they do not quantify the skill of the surgeon in terms of dexterity parameters such as accuracy, time and economy of movement. There have been attempts made to measure similar parameters, but these involve video watching and thus depend on the subjectivity of the viewer (Den Boer et al 1999). Motion analysis systems can measure hand movements they could also be used for the assessment of surgical dexterity.

ICSAD (Imperial College Surgical Assessment Device) – Generic Technical Skills

This system consists of a commercially available electromagnetic tracking system (Isotrak II, Polhemus Inc, USA), which is connected to a portable computer through a standard RS-232 (serial) port. The Isotrak II system (Picture 2.1) consists of an electromagnetic field generator and two sensors which are attached to the dorsum of the surgeon's hands at a standardised position. Bespoke software is used for data retrieval and analysis. This uses a software called 'Track 3D' to collect positional data and a software called 'Track' for converting this data to dexterity measures such as the number and speed of hand movements, the distance travelled by the hands and the time taken for the task (Picture 2.2). Recently developed new software, ROVIMAS (Robotic video motion analysis software), which allows the stream video files along with the dexterity data in order to 'zoom' into certain key steps of a procedure and filter out data from moments where the procedure is eventless such as while waiting for instruments or delays due to equipment failure (Dosis¹ et al 2004).

Studies have demonstrated the construct validity of ICSAD with respect of a range of laparoscopic tasks (Datta¹ et al 2001). There is a strong correlation between dexterity and previous laparoscopic experience on a simple task in a box trainer (Taffinder³ et al 1999),

and on more complex tasks such as laparoscopic cholecystectomy on a porcine model (Smith² et al 2002). Experienced and skilled laparoscopic surgeons are more economical in terms of the number of movement and more accurate in terms of target localisation and thus use significantly lower path lengths, which is a measure of accuracy. Datta¹ et al 2001 have also demonstrated the construct validity of motion analysis on open surgical bench model tasks. One study has shown a strong correlation between motion analysis and qualitative assessment using the likert rating scale in OSATS (Datta⁴ et al 2002). However once again ICSAD measures generic technical skills which usually improve with number of procedures, but does not assess the quality of the operation or task performed.

Advanced Dundee Endoscopic Psychomotor Trainer – Generic Technical Skills

Motion tracking can be based on electromagnetic, mechanical or optical systems. The Advanced Dundee Endoscopic Psychomotor Trainer (ADEPT), another motion tracking system was originally designed as a tool for the selection of trainees for endoscopic surgery, based on the ability of psychomotor tests to predict innate ability to perform relevant tasks (Hanna¹ et al 1997). Studies have demonstrated the validity and reliability of ADEPT (Francis^{1,2} et al 2001). Optical motion tracking systems depend on infrared sensors which are attached to the surgeon's arms (Emam et al 2002). These systems consist of infrared cameras surrounded by infrared light emitting diodes. The infrared light is reflected of markers which are placed on the limb of a surgeon. Software is used to extrapolate the positional data of the markers to data on movement analysis. The disadvantages of such optical systems are that they suffer from line of vision disturbances. If the camera and signal generators become obscured from the markers, the resulting loss in link leads to lost data. To overcome this shortcoming more than a few cameras and markers are required. Signal overlap also prevents the use of markers on both limbs making these systems more restrictive.

Virtual Reality – Generic Technical Skills

Virtual reality simulators use motion analysis, measuring path length of the surgeon's hand.

An account of virtual reality simulators is given in the section below.

Table 2.8 – Dexterity Analysis Systems Papers (with 20 or More Subjects)

Number of Papers	Total Number of Subjects	Conclusions
28	1454	Motion analysis is able to discriminate experts from novices using pathlength. It is not a marker of quality of the operation. This generic skill improves with training, and can be affected by stress e.g. sleep deprivation.

Virtual Reality – Generic & Specific Technical Skills

Table 2.9 summarises the 23 virtual reality (VR) simulator papers with more than 20 subjects. VR simulators have face, content, predictive and construct validities. They are reliable and can assess generic technical skills (time and motion). Newer VR simulators with better haptics can assess specific technical skills and errors.

Virtual reality is defined as a collection of technologies that allow people to interact efficiently with three dimensional computerised databases in real time by using their natural senses and skills (McCloy & Stone 2001). Surgical virtual reality systems allow interaction to occur through an interface, such as a laparoscopic or endoscopic frame with modified instruments.

The minimally invasive surgical trainer-virtual reality (MIST-VR) system was one of the first virtual reality laparoscopic simulators developed as a task trainer. The system was developed as the result of collaboration between surgeons and psychologists who performed a task analysis of laparoscopic cholecystectomy. This resulted in a toolkit of skills needed

to perform the procedure successfully (McCloy & Stone 2001). These were then replicated in the virtual domain by producing three dimensional images of shapes, which users can manipulate.

As virtual reality simulators are computer based systems they generate output data, or what is commonly referred to as metrics. In an international workshop a group of experts reviewed all methods of assessment and suggested parameters that should constitute output metrics for the assessment of technical skills (Satava ¹ et al 2003), included were variables such as economy of movement, length of path, and instrument errors. The MIST-VR system has been validated extensively for the assessment of basic laparoscopic skills (Taffinder ¹ et al 1998). Grantcharov and colleagues over the last five years with their studies on laparoscopic and endoscopic simulators have shown the benefit of training on a virtual reality simulator (Adamsen et al 2005, Everbusch et al 2004, Grantcharov ^{1,3,4} et al 2001,2003,2004).and the ability of detecting surgeons under stress (night on-call) generic technical errors (Grantcharov ¹ et al 2001). The latter was consistent with Taffinder ² et al 1998 findings in their research letter assessing junior surgeons' technical skills after sleep deprivation.

In the last few years there have been developments in other virtual reality simulators e.g. Reachin[®], ProMis[®], Symbionix[®] etc, these systems are of higher fidelity and have tactile feedback (haptics). In the future, there is a possibility that higher fidelity virtual reality systems may be used as procedural trainers and for the assessment of procedural skill. One of the main advantages of virtual reality systems, in comparison to dexterity analysis systems, is that they provide real time feedback about skill based errors.

Table 2.9 – Virtual Reality Simulators Papers (with 20 or More Subjects)

Number of Papers	Total Number of Subjects	Conclusions
23	1196	VR simulators have face, content, predictive and construct validities. They are reliable and can assess generic technical skills (time and motion). Newer VR simulators with better haptics can assess specific technical skills and errors. They can be used to assess the affect of stress on surgical performance.

Quality of the Final Product – Specific Technical Skills

As outcomes after surgery is difficult to depend solely to surgical technique and as adverse outcomes from poor technique may not be apparent for many years such as recurrence after cancer resections, some researchers have suggested the idea of using outcome measures after performance on bench models. Szalay et al assessed the ‘quality of the final product’ after performance on six different bench models (Szalay et al 2000). This outcome analysis consisted of a four item rating scale on a 5 point likert scale which assessed completeness, aesthetics, function and overall quality of performance. This was performed by senior surgeons who were blinded to the skill of the trainee. The investigators found that the method possessed construct validity. They demonstrated a correlation between OSATS and the final product assessment suggesting that analysis of the final product may surmount some of the problems involved with live ratings. However, the inter-rater reliability between the assessors who performed the outcome analysis was lower than that for OSATS (Regher et al 1998).

Datta ⁴ et al 2002 assessed the leak rates after performance of a vascular anastomosis on a bench model and found that there was significant correlation between the leak rate and surgical dexterity. Hanna et al studied that quality of knots performed laparoscopically

using a tensiometer and derived a knot quality score as an index of knot reliability (Hanna² et al 1997). However there is no tool to date to assess the quality of a live operation in animals or humans. This would be advantageous as technical skills (non-patient factor) could be added and compared to patient factors which affect outcome.

Error Scoring – Specific Technical Skills

Table 2.10 summarises all the 27 papers with more than 20 subjects using error scoring. Error scoring can assess generic and specific technical skills. It demonstrates construct validity and reliability. Feedback on technical skills errors improves future surgical performance. Under stress e.g. sleep deprivation novice surgeons can make more technical skills errors.

One way of measuring technical ability may be to examine performance for the number of errors executed during a procedure. One study assessed performance while trainees performed a small bowel anastomosis on a porcine model and generated a cumulative error score (Steele et al 1992). Eubanks et al developed a minor and major error score of the technical errors executed during a laparoscopic cholecystectomy in addition to a structured checklist. Such errors were weighted according to their severity (Eubanks et al 1999). Using a categorisation of errors based on the external error modes framework, Joice² et al 1998 identified errors executed during the performance of the laparoscopic cholecystectomy. This categorisation of errors included procedural errors which were done either because of the omission or the re-arrangement of steps during the procedure or execution errors which were failure of the surgeon to perform a step correctly (Joice² et al 1998). Seymour et al developed an error score for the purposes of a study evaluating the transfer of skill from virtual reality to real procedures. They found that those trained on the

VR simulator (MIST-VR, Mentice, Sweden) performed 5 times fewer errors as compared to those who received conventional training (Seymour ¹ et al 2002).

Bann ⁵ et al 2005 recently demonstrated that assessing simple technical skills errors was a predictor of technical skills and performance on bench model tasks. What is yet to be answered is whether the use of such models and principles can shorten the qualitative surgical learning curve.

However the situation is more complex as one study demonstrated a large number of minor technical skills errors performed by expert surgeons in laparoscopic surgery, but few significant major technical skills errors (Sarker SK ³ et al 2005). The experts seem to migrate to a high level of minor technical skills errors but then migrate back when essential aspects of the operation are being done, and therefore executing few major technical skills errors.

Table 2.10 – Error Scoring Papers (with 20 or More Subjects)

Number of Papers	Total Number of Subjects	Conclusions
27	1296	Error scoring can assess generic and specific technical skills. It has construct validity and reliability. Feedback on technical skills errors improves future surgical performance. Under stress e.g. sleep deprivation novice surgeons make more technical skills errors. Error scoring can be used in VR simulators, Open and MIS.

2.15 Conclusions

The structured assessment of generic and specific technical skills in surgery can parallel the structured curriculum based training of surgical skills. Figure 2.1 is a modified recommended format for such a training-assessment cycle which should ensure that progression from one level to the higher one is based on robust criteria, originally published 2 years ago (Moorthy² et al 2003). The modifications enable the programme of assessment to be potentially practical and feasible. Such a shift in surgical learning will also ensure that trainees get feedback all throughout their training programmes and get opportunities to address deficiencies in performance in a continuous manner.

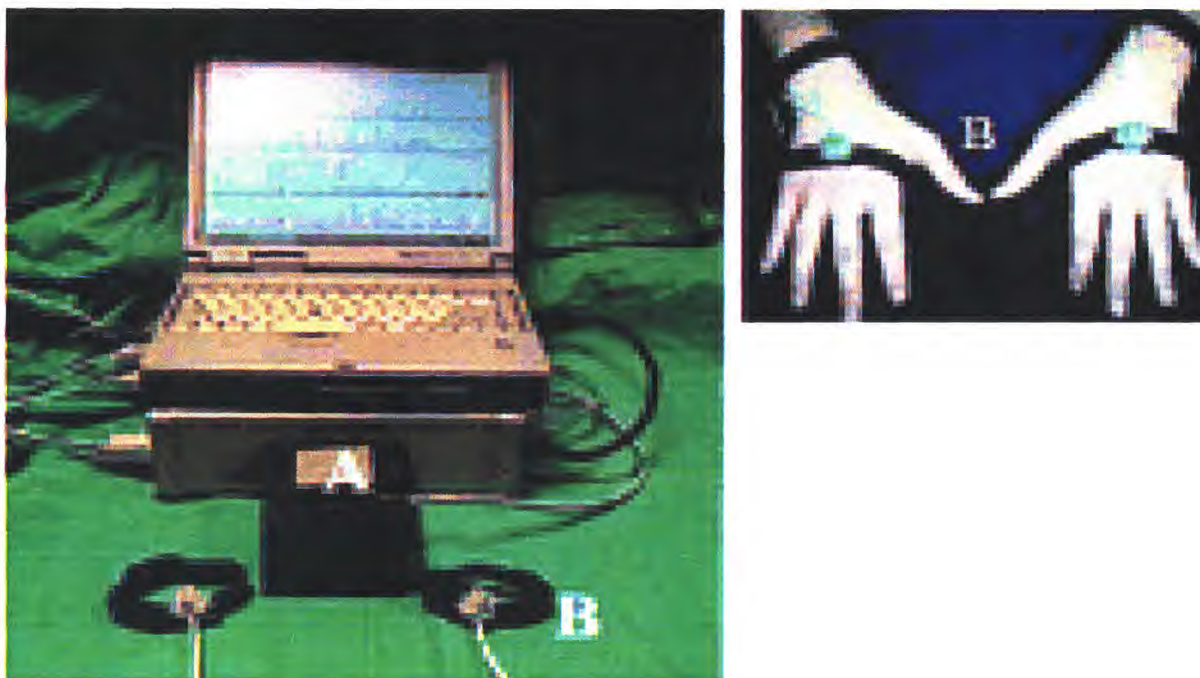
End of course evaluations should be made structured, especially on those such as the basic surgical skills course where success and not merely attendance is important. This would also give course participants an opportunity to gain an insight into the utility of the course and allow centres that run such courses to strive to improve the quality of teaching ensure standardisation and change course formats according to the performance of participants. Teaching of technical skills in surgery is important as how the teacher demonstrates or supervises the trainee may affect the trainee's acquisition of technical skills (Sarker SK² et al 2005).

In the light of recent public and media attention, senior surgeons or senior trainees need to be methodically assessed. There have been several cases of very experienced surgeons performing complex procedures for several years with poor patient outcome. Further research should focus on the predictive validity of these assessment measures and more importantly, it will be essential to demonstrate that generic and specific technical skills assessments correlate with patient outcome measures. This will need the development of a large database of surgeons before these assessment measures can be employed for

suggesting alternate careers to consistently poorly performing surgeons and for purposes of revalidation.

The past decade has seen considerable progress in the development of structured methods of assessment of technical skills, both generic and specific. This has been made possible in spite of the initially perceived difficulty. The surgical community now has a wide range of methods of assessment available which have been validated. This has been further enhanced with advancing technological innovations in assessing technical skills. These are now ready to be used initially during training to make feedback more objective, make progression criteria based, and help poorly performing trainees take remedial actions and as research methods evaluating training interventions. However, it must be emphasized that technical skills are only a part of surgeon's competence and there is a need to integrate the assessment of technical skills with decision making and team performance in order to develop methods that comprehensively assess surgical competence.

Picture 2.1 – Imperial College Surgical Assessment Device (ICSAD)



Picture 2.2 – ICSAD Tracing

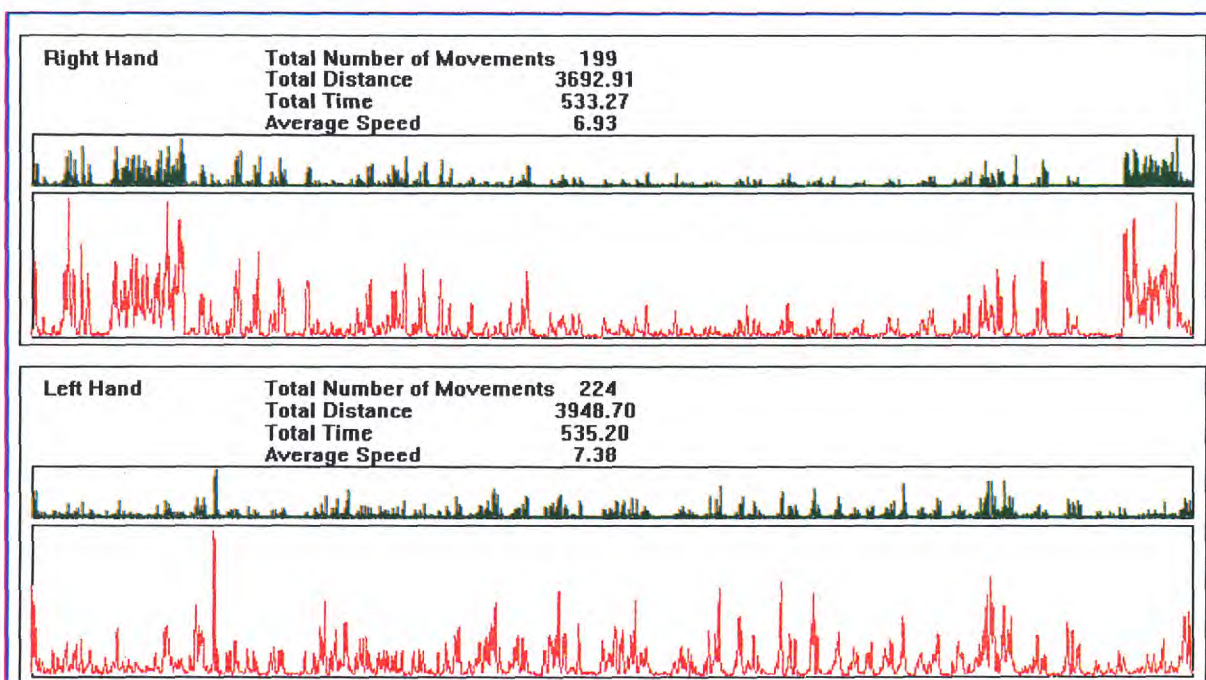
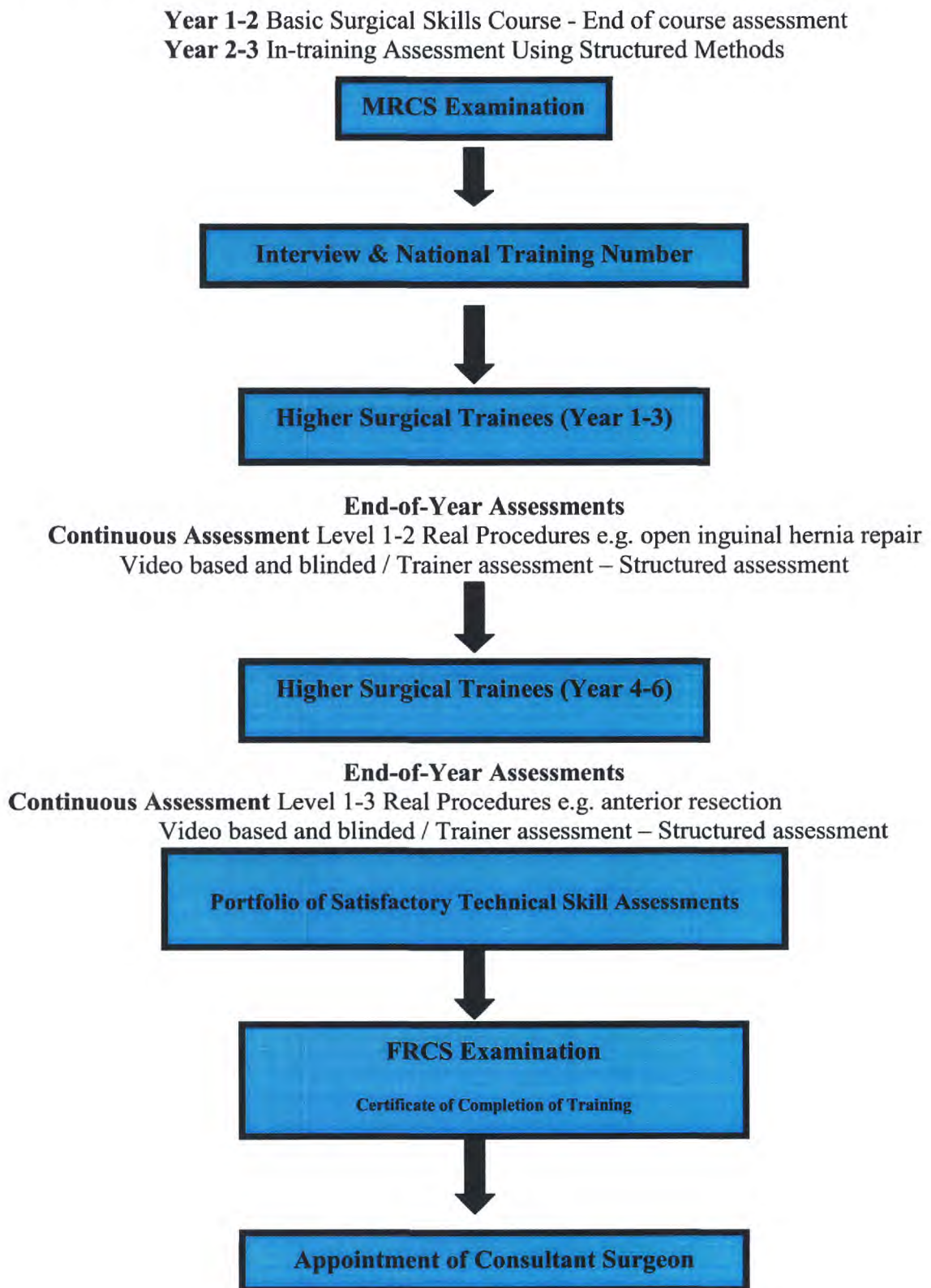


Figure 2.1 - Assessment of Technical Skill in Surgical Training



Chapter 3

Human Error & Reliability

3.10 Overview

Making errors is part of normal human behaviour (Reason ¹ 1990). However when errors have significant consequences, or occur in high risk industries they become of paramount importance. Society and the media are generally intolerant of people making errors which may cause human suffering, and therefore cultivate a blame culture. As a result of high profile accidents which have caused many deaths, high risk industries like aviation, space travel, military, nuclear and oilrigs (Barach & Small 2000) have spent the last few years investing in the understanding, identification and error prevention training in their respective industries. All of these industries have the financial resources to invest in the research and development of error prevention. Most of these industries work in controlled environments and hence a lot of their error training and prevention have concentrated on simulations and re-enacting major incidents (Testoni et al 2004).

Medical errors are periodically brandished across the newspapers and the news, which always attract the public's attention and concerns. For the media medical errors equate to patient harm, but this is not always the case. There has been little research in why and how errors occur in the healthcare industry. Most of the current research has been on the prevention of drug errors and the estimation and epidemiology of adverse events occurring to inpatients (Brennan et al 1991, Leape ¹ et al 1991, Leape ² 2000, Thomas ¹ et al 2000). However these epidemiological estimates are only the tip of the error iceberg in healthcare.

In particular, surgery as a high risk specialty, has been shown to have a detrimental effect on patients with one study quoting an annual incidence rate of adverse events among patients having an operation as 3.0%, of which half were preventable (Thomas ² & Brennan 2001). Almost all surgeons and healthcare professionals are well motivated and care about patients' treatment and safety (Carter 2003).

Eric Hollnagel (1998) points out that the term error has historically been used in three different senses: as a cause of something, as the event itself (the action) or as the outcome of an action. These three senses all appear in the media and the medical literature (Table 3.0). The distinctions are not absolute in that many uses of the term involve both cause and consequence to different degrees, but they do have a very different emphasis. For instance, the UK National Patient Safety Agency has found that patients equate 'medical error' with a preventable adverse outcome for the patient. Terms like 'adverse event', although technically much clearer, just seem like an evasion or a way of masking the fact that someone was responsible. The term 'medical error' cannot, however, be used simply to refer to adverse outcomes for patients, if only because numerous errors occurs which never lead to harm (Vincent ² 2006).

Table 3.0 - Errors

Different Senses of the Word Error	
<i>As a cause</i>	Rail crash due to 'human error' Error by surgeon caused death of patient
<i>As an event or action</i>	I forgot to check the tire pressure I picked up the wrong drug ampoule
<i>As a consequence or outcome</i>	Salt in my coffee 98,000 medical errors each year in U.S.A.

The most precise definition of error, and most in accord with everyday usage, is one that ties it to observable behaviours and actions. As a working definition, John Senders and Neville Moray (1991) proposed that an error means that something has been done which:

- Was not desired by a set of rules or by an external observer
- Led the task or system outside acceptable limits
- Was not intended by the person performing the action

This definition of error, and other similar ones (Hollnagel 1998), imply a set of criteria for defining an error. First, there must be a set of rules or standards, either explicitly defined or at least implied and accepted in that environment; second, there must be some kind of failure or 'performance shortfall'; third, the person involved did not intend this and must, at least potentially, have been able to act in a different way. All three of these criteria can be challenged, or at least prove difficult to pin down in practice. Much clinical medicine for instance is inherently uncertain and there are frequently no

guidelines or protocols to guide treatment. In addition the failure is not necessarily easy to identify; it is certainly not always clear, at least at the time, when a diagnosis is wrong or when at what point blood levels of a drug become dangerously high. Finally, the notion of intention, and in theory at least being able to act differently, is challenged by the fact that people's behaviour is often influenced by factors, such as fatigue or peer pressure, which they may not be aware of and have little control over. So, while the working definition is reasonable, we should be aware of its limitations and the difficulties of applying it in practice.

In this present chapter there will be an outline of human error and its role in Surgery:

1/ Human Error – Definitions, Types & Classifications

2/ Human Error Identification Assessment Techniques

3/ Errors in Healthcare

4/ Assessing Errors in Surgery

This chapter therefore examines the principles and definitions of error, its detection and assessment, techniques applicable to surgery. The later of which we will then apply and adapt so we can assess the technical skills performed during open and laparoscopic operations which are subject to errors.

3.11 Human Error – Definitions, Types & Classifications

Errors as a result of human actions and behaviour are an unavoidable occurrence, and is part of normal human behaviour. This section will discuss and define human error in view of cognitive psychology.

Correct performance and systematic errors are symbiotic. Broadly based analysis of recurrent error is essential to understand the processes of human thought and action in relation to the occurrence of errors.

Errors appear in very similar forms across a range of human mental activities. It is possible to identify comparable error forms in action, speech, perception, recall, recognition, judgement, problem solving, decision making, and concept formation.

Error – Definitions

Error – generic term to encompass all those occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of chance.

Slips and Lapses – are errors which result from some failure in the execution and/or storage stage of an action sequence, regardless of whether or not the plan which guided them was adequate to achieve its objective.

Mistakes – are deficiencies or failures in the judgemental and/or inferential processes involved in the selection of an objective or in the specification of the means to achieve it, irrespective of whether or not the actions directed by this decision-scheme run according to plan.

Further subtype definitions of errors and actions are summarised in the Appendix Chapter 3.

Error - Types

This term relates to the presumed origin of an error within the stages involved in conceiving and then carrying out an action sequence. These stages can be described under three broad headings: planning, storage and execution.

<i>Cognitive Stage</i>	<i>Primary Error Type</i>
Planning	Mistakes (Failures of Expertise/Lack of Expertise)
Storage	Lapses
Execution	Slips

Error – Forms

These are recurrent varieties of fallibility that appear in all kinds of cognitive activity, irrespective of error type. Error forms are so widespread that it is extremely unlikely that their occurrence is linked to the failure of any single cognitive entity.

At the ***skill*** level, actions are automatic, and are enacted by way of “stored patterns of pre-programmed instructions” (Rasmussen 1983). These actions are frequently performed and are often said to ‘come naturally’ to the operator. For example performing a routine elective inguinal hernia repair

On a ***rule*** level, tasks are completed using stored sets of rules. These rules consist of familiar, rehearsed patterns of actions. Tasks which use rule-based cognitive mechanisms require a greater degree of thought than skill-based tasks, as the rules which need to be applied to complete the situation must be selected. For example ensuring swabs and instruments are correct after completion of an operation.

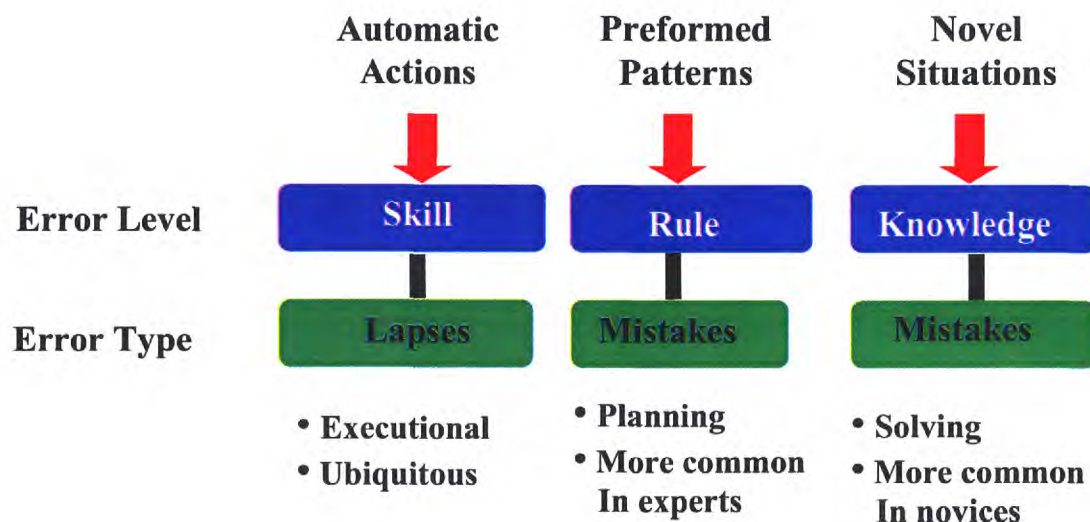
On a ***knowledge*** level, unfamiliar tasks are performed with a high degree of conscious thought as the operator attempts to devise a novel solution to a situation which has not previously been encountered. For example a junior surgeon performing a laparotomy unsupervised for the first time.

Reason used Rasmussen’s classification as a framework for his categorization of errors (Figure 3.1), attaching a specific type of error to the three levels of cognitive

performance: errors in execution of skill-based tasks were termed *lapses*, and errors in execution of rule and knowledge based behaviours were termed *mistakes* (Reason 1990).

Lapses characteristically relate to an error in the actual execution of the task, whilst mistakes are more abstract errors, relating to errors in planning (a strategy not suitable for the situation is carried out), or in problem-solving. In the case of lapses, the plan is correct but the actions are carried out incorrectly e.g. not correctly placing a clip on the cystic duct, leading to an error; whereas in the case of mistakes, the actions are carried out correctly, but instead the plan is incorrect e.g. performing a laparotomy thinking it was a ruptured abdominal aneurysm but was actually a severe heart attack.

Figure 3.1 - Rasmussen's Error Level Model & Reason's Error Type Classification



Error - Surgical Definitions

It is essential to examine errors in the context of a surgeon performing an operation on a patient; this will explore the issue of error and accountability. During an operation, there can be a specific action that is the error e.g. damaging the femoral vein. This is a specific event and is called the *coal-face error (active)*. The importance of this specific definition is that it provides a reference point in defining surgical errors. The other type of error is the *systemic error (latent)*, which is a series of events that led up to the error, sometimes referred to as the root cause e.g. theatre list starting late, patient not properly prepared, wrong surgical instruments etc. These surgical definitions (Cuschieri ² 2005) correspond roughly to Reason's terminology of active and latent errors. Therefore there are two types of surgical error regarding an operation or procedure:

1/ Coal-Face (Active) - a specific event resulting in an error

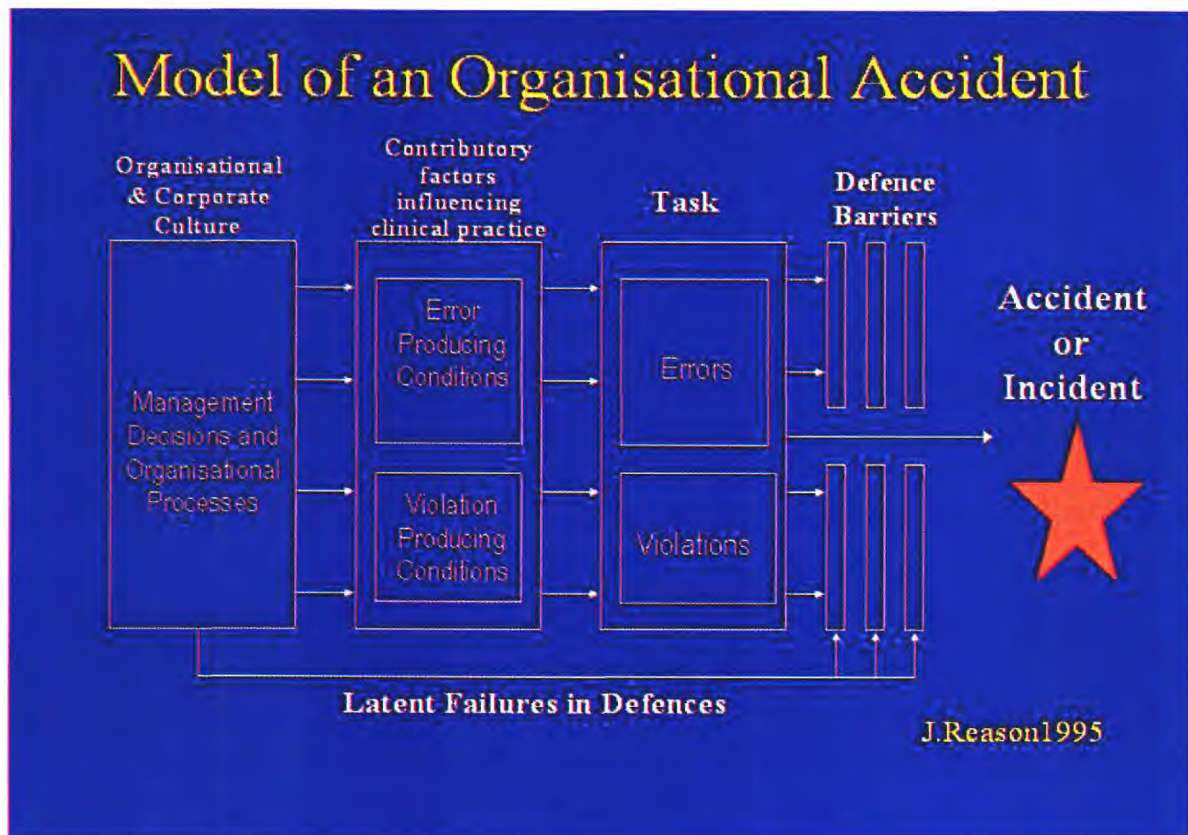
For example the surgeon is performing a routine technically easy laparoscopic cholecystectomy but cuts the common bile duct. The coal-face error is influenced by factors directly affecting the surgeon:

<i>Physiological Factors</i>	e.g. fatigue, sleep, hunger, state of health
<i>Emotional Factors</i>	e.g. mood states (e.g. anxiety), life events, confidence, motivation, personality
<i>Cognitive Factor</i>	e.g. thought processes involved in tasks such as planning and decision-making (cognitive biases)
<i>Stress Factors</i>	e.g. difficult or novel tasks, work load, interruptions, urgent tasks etc

2/ *Systemic (Latent)* - a series of events resulting in an error

For example patient arrives late on day surgery unit for a left sided hernia repair, patient is not skin marked for inguinal hernia repair, the operation list has the wrong side of operation, and the surgeon performs the operation on the wrong side. Figure 3.2 summarizes Reason's model of latent failures.

Figure 3.2 - Reason's Model of Latent Failures



Error - Surgical Classification

Lucien Leape (Brennan et al 1991) has further aided the understanding of the nature of surgical error by classifying it (Table 3.1). This classification brings errors directly into the realm of clinical surgical practice. Surgical care in the *pre-operative* period is principally that of *diagnosis*, of obtaining the appropriate amount of information (from history, physical, laboratory and imaging studies, etc.), a decision *to operate or not*, and making the *appropriate preparations* for the surgical operation or procedure. During the *intra-operative* stage of the surgical operation or procedure, the *correct therapy or technique* must be performed with competent cognitive (tasks & sub-tasks, knowledge & recognition of specific anatomy, dynamic intra-operative decisions & actions) and psychomotor (generic and specific technical skills) proficiency. Finally the *post-operative* period is a time of *vigilance and monitoring* to ensure that there are no complications due to treatment after the surgical procedure. There are also contributing factors that can happen during any of the above phases, such as failure of team co-operation or communication, or equipment or instrument failure.

Table 3.1 – Error: Surgical Classifications

Category	Failure & Error
Diagnostic (errors or delay in diagnosis)	Failure to Employ Indicated Tests Use of Outmoded Tests or Therapy Failure to Act on Monitoring or Test Results
Treatment (error in performance of an operation, procedure, or test)	Error in Administering Treatment Error in Dose or Method of Using a Drug Avoidable Delay in Treatment or Response to Abnormal Test Inappropriate Care
Preventive (failure to provide prophylactic treatment)	Inadequate Monitoring or Follow-up
Other	Failure of Communication Equipment Failure

3.12 Human Error Identification Assessment Techniques

Human Reliability Assessment (HRA)

Human reliability assessment provides a structured approach for understanding how human performance contributes to risk.

A number of industries have embraced HRA as a solution to their human factors and safety problems or have been required to apply them due to public or governmental pressure. The nuclear industry was the first to develop and apply human reliability assessment as a field in its own right (Kirwan ² 1994). The public fear of the risk of a nuclear reaction and the responsibility placed in the hands of a single control room operator were ensured that both the human and technological possibilities of error and

breakdown were subjected to intense scrutiny. Other industries have also adopted HRA as a risk assessment strategy as the reliance on retrospective accident analysis would not preclude an incident occurring, just the opportunity to prevent a similar event from re-occurring in the future.

Since then, HRA has been applied in many “high-risk” industries including aviation and aerospace, rail, shipping, air traffic control, automobile, offshore oil and gas, chemical, and all parts of the military (Humphreys 1995). In total there are 85 HRA techniques (Lyons et al 2004). However there are only 35 primary HRA techniques, which have a practical application in healthcare or which are well established elsewhere and had potential application (Lyons et al 2004). Of these only some of these techniques are practically applicable to healthcare and surgery.

The key steps for human reliability assessment are (Kirwan ² 1994):

1. Understand and describe the tasks which people carry out

It is necessary to understand what people should do before we can start to predict how humans can impact on the performance of a task. This is usually developed out of standard operating procedures, but will also include insights on the tasks from those who actually carry them out. This is done by outlining a task analysis for the procedure.

2. Predicting what errors can occur

For each task step identified, the human errors that can occur are considered. A basic assessment might consider errors such as omitting the task step, not carrying it out at the right time or carrying it out inadequately. Some errors may not be plausible and can be ruled out at this stage.

3. Likelihood and consequences of error

For plausible errors, it is necessary to consider how likely they are to occur and what would be the consequences if they did occur.

Kirwan ¹ & Ainsworth 1992 grouped the techniques in Table 3.2 into five categories spanning the principal types and purpose of HRA analysis. Some techniques are primarily descriptive or concern basic data gathering (Table 3.2). These are often used as a prelude to more sophisticated approaches involving simulation, human error analysis and human error quantification. Techniques may be used separately, but more often in combination.

Table 3.2 - Range & Scope of HRA Techniques

Type of Technique	Description
Data Collection	Collection of information on incidents, goals, tasks
Task Description	Taking the data collected and portraying this in a useful form
Task Simulation	Simulating the task as described and changing aspects of it to identify problems
Human Error Identification and Analysis	Uses task description, simulation and/or contextual factors to identify the potential errors
Human Error Quantification	Estimated the probability of the errors identified

3.13 Assessing Errors in Healthcare

There are important differences between healthcare and other industries. Firstly, healthcare consists of an extraordinarily diverse set of activities. Healthcare encompasses the mostly routine, but sometimes highly unpredictable and potentially harmful activities e.g. surgery, trauma, cardiac arrest.

Unlike other industries, the healthcare system relies on the human-human interaction as opposed to human-machine interaction, and therefore prone to human factors which can cause errors. There is no central focus of the healthcare staff's work such as in an aircraft or on an oil platform. In contrast, the work is focused on a single patient as part of a flow of large numbers of patients. None of this is to say that HRA techniques should not be applied and utilized in healthcare.

3.14 Assessing Errors in Surgery

Human Reliability Assessment Relevant to Technical Skills Errors

In the following section we will concentrate on HRA techniques which are applicable to Surgery and then mainly concentrate on those which can directly be used in assessing technical skills errors in Surgery. Table 3.2 summarizes the range and scope of HRA techniques. Of these the following are applicable to Surgery: Data Collection, Task Description, Task Simulation, and Human Error Identification & Analysis.

Data Collection

Collection of information on adverse events and critical incidents occur in all hospitals in the UK. However the majority of cases are unreported. Another route where surgical data is collected and presented is the monthly mortality and morbidity meetings by the surgical units in each hospital. These compulsory audits may highlight errors or failings in the pre-operative, intra-operative and post-operative phases of patient care.

Task Description

Auditing surgical data over a period of time e.g. one or 5 years may be beneficial in identifying errors or discrepancies in patient care. An increase in mortality by an individual surgeon may be identified in this way. All surgeons must present their mortality and morbidity data on a yearly basis. The Association of Cardiac Surgeons collects data, as a result of the Bristol inquiry, from all cardiac surgeons in the UK, so if increased mortality rates are identified they can be acted upon.

Task Simulation

Simulating the task as described and changing aspects of it to identify problems can be achieved. This can be done in the clinical skills laboratory e.g. setting up bench models for individual tasks for individual operations. Also in the last few years laparoscopic simulators have increased in their fidelity, enabling trainees to learn in a controlled environment about technical skills errors – generic and specific.

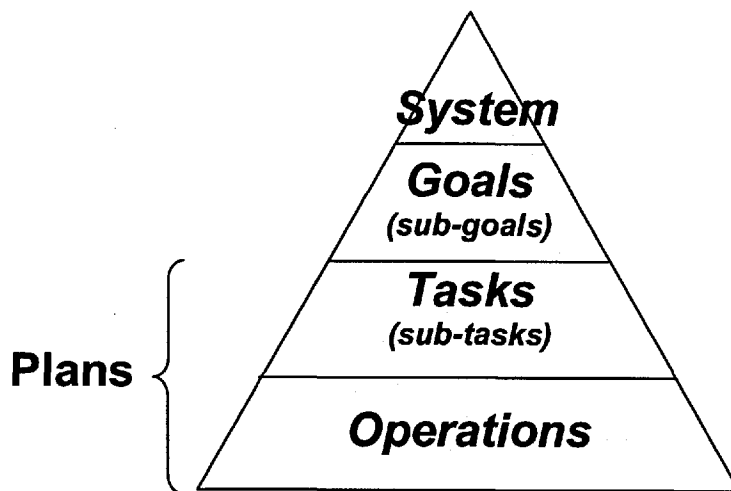
Human Error Identification & Analysis

This uses task description, simulation and/or contextual factors to identify the potential errors. As a result this method is the most applicable to surgery, and assessing technical skills error. There are two specialized methods within this group which can be used in surgery: task analysis and hierarchical task analysis. The following are accounts of these specialized methods.

Task Analysis

Task analysis is a tool employed in analysis of systems whose complexity prevents straightforward analysis (Embrey 2000). Task analysis employs a number of terms with specific meanings (Fig.3.3)

Figure 3.3 - Terms Used in Task Analysis



Analysis takes place on defined *systems* (e.g. laparoscopic cholecystectomy)

Systems have a *goal*, which is the aim of working on the system (e.g. safe removal of a patient's gall bladder)

Goals can be broken down into *sub-goals*, which more clearly demonstrate the aims required to complete the sub-goal.

Goals are achieved by multiple *tasks*, which describe how the goal is achieved (e.g. dissect the gall bladder from the liver).

Tasks can also be further broken down into *sub-tasks*.

Tasks are described by *operations*, which physically detail how to perform a task (e.g. dissect away the tissue connecting the gall bladder to the liver)

Plans dictate the sequence and conditions in which *operations* are to be performed. (e.g. prepare the abdomen before draping the patient).

Task analysis breaks down complex *tasks* into their component *sub-tasks* that must be performed for the *task* to be completed, to achieve the *goals* and *sub-goals* of the *system*.

Sub-tasks can, of course, be further analyzed to assess by what *operations* they are achieved by (Kirwan ² 1994; Johnson 1992; Shepherd 1989). Task analysis can seem an endless undertaking, decomposing tasks into simpler and simpler sub-tasks. However, it

can be tailored to a required level of expertise, where either the operator in the system is deemed competent to complete the sub-task, or an analyst of the system can understand the task, without further redefinition.

The ' $P \times C$ rule' can also be applied to cessation of task analysis. This refers to P , the *probability* of the operator failing to complete the task, and C , the *cost* (personal, financial, medical) of not completing this task. If $P \times C$ is high (i.e. there is a high chance of the operator failing to complete the task, or the cost of non-completion is very high), then a further set of sub-tasks is required to better define a task (Kirwan ² 1994). In surgery such an example of a high $P \times C$, would be an inexperienced surgeon trying to repair a torn inferior vena cava in a trauma patient who is severely bleeding.

Task analysis allows human actions to be examined in a more discrete manner, allowing easier and more effective analysis. The nature of task analysis endows it with two points which make it a very powerful tool: it employs an objective simplification process, which ensures that for a given system, a standard set of sub-tasks and operations should exist, regardless of who performs the task analysis (though depth of analysis may vary between analysts or operators); and its decomposition of complex tasks into simple steps allows non-experts to analyze systems (Embrey 2000).

Therefore task analysis can be used to break down a surgical procedure into goals, tasks and operations. However task analysis does not incorporate all the aspects of completing the procedure e.g. asking the anaesthetist if the surgeon can start the operation. Therefore

we need hierarchical task analysis to incorporate all the salient parts in successfully completing a surgical procedure.

Hierarchical Task Analysis

Hierarchical task analysis (HTA) is a branch of task analysis where an emphasis is placed upon the organization within the system on the completion of a task or procedure. HTA exists as a combination of tasks, their constitutive operations, and plans, all describing how to achieve their antecedent sub-goal, in terms of chronological completion of tasks.

The *hierarchy* stems from the plans, which control when the operations are to be performed to complete each task. Because each individual sub-goal has its own associated hierarchy of operations, the operative description for completion of the tasks leading to a sub-goal can be customized to the level of description that is required for the operator to carry out the task.

This method of analyzing human actions is particularly useful as it maintains a link between the goal of the system, and the actual actions of the operator that realize that goal. HTA is often represented by form of tables or flow charts, which reflect the chronological nature of this method of analysis. It is a very strong method of analyzing inflexible systems, due to its very clear breakdown of progression of operations through a system. However, once a degree of uncertainty is introduced into a system, HTA can quickly become unmanageable. HTA is thus primarily reserved for analyzing systems which vary very little from a set framework (Kirwan ² 1994).

Task analysis has been used in surgery in the aim to identify and quantify errors executed in performing an operation, Table 3.5 summaries all the papers on surgical task and error analysis.

Cao ¹ et al 1996 (Table 3.3) in their paper was able to break down the movements and tasks/subtasks into generic and specific technical skills using data from laparoscopic training workshops. Joice ² et al 1998 (Table 3.4) further advanced this technique and used real time laparoscopic cholecystecomies to further quantify the tasks and sub-tasks performed in this procedure. Cao ² et al 1999 (Figure 3.4) incorporated that a surgeon must interact with the surroundings and key personnel to complete the operation. In their study they used two different settings to encapsulate HTA in surgery: laparoscopic training workshop (live pigs) and the performance of laparoscopic surgery in a live operating theatre (live human). They chose three common laparoscopic operations: laparoscopic cholecystectomy, laparoscopic trans-abdominal pre-peritoneal inguinal hernia repair and laparoscopic fundoplication. They used an observational technique to assess the hand movements of the surgeons, tasks/subtasks the surgeons perform and how, what and who with the surgeon interact with in completing and performing the operation. They used videotaping to observe the outside environment (outside patient's abdomen) and the inside environment (inside patient's abdomen). Den Boer et al 2002, Mehta et al 2002 have further categorised the process of laparoscopic surgery.

Eubanks et al 1999 and Tang ¹ et al 2004 have used task analysis as a basis to assess technical errors performed by surgeons whilst performing live human laparoscopic

cholecystectomies. Eubanks based his scoring system on 30 laparoscopic cholecystectomies and Tang based his Observational Clinical Human Reliability Assessment (OCHRA) scoring on 200 laparoscopic cholecystectomies. There have been further developments recently using task analysis and HTA in the teaching and appraisal environment. Velmahos et al 2004 and Tang² et al 2005 have used task analysis to assess teaching and the performance of course participants in laparoscopic surgical skills courses.

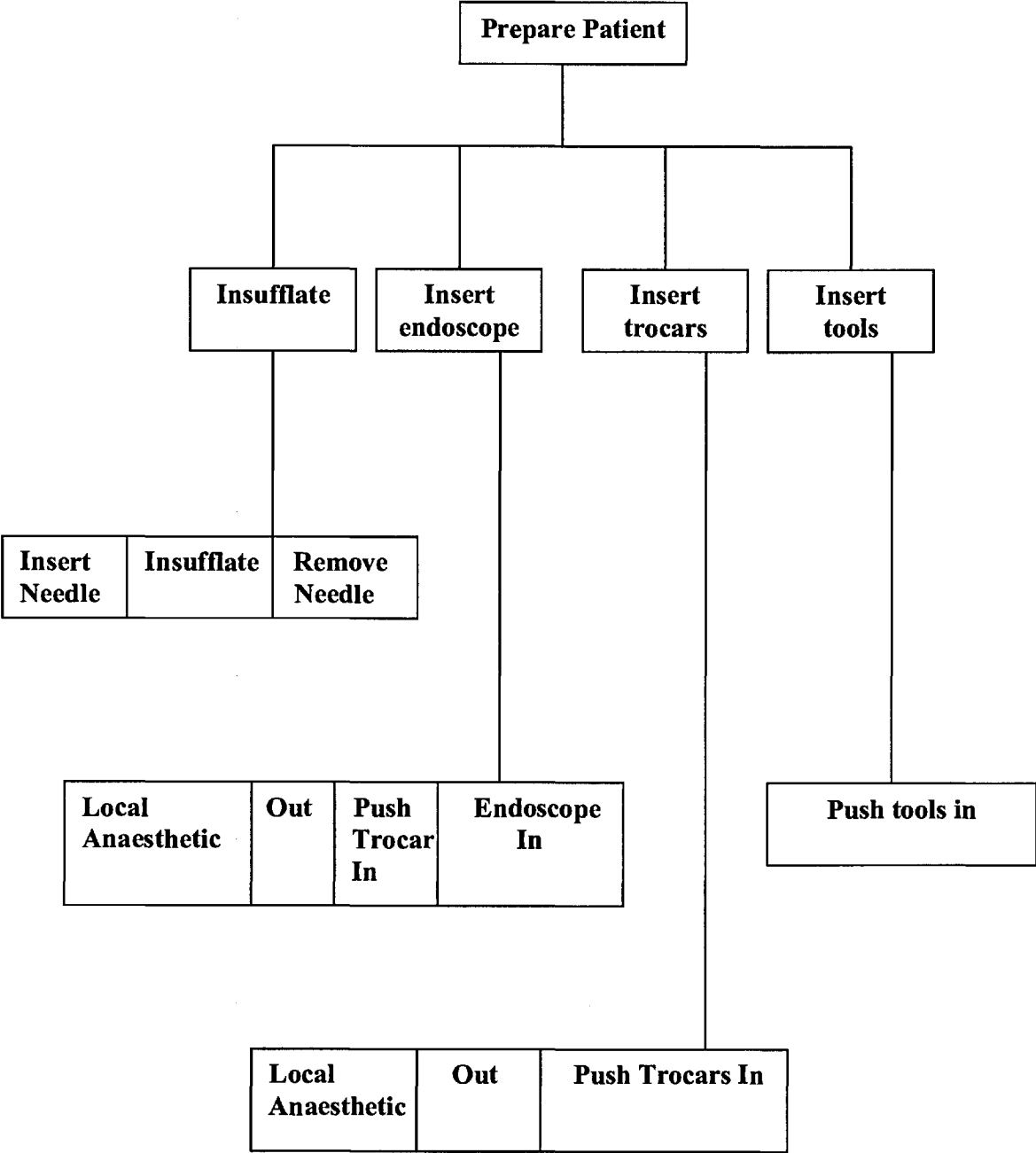
**Table 3.3 - Task Analysis of Laparoscopic Tissue Dissection
Generic Technical Skills (Cao et al 1996)**

Task	Subtasks	Tool	Begin	End
Dissect Tissue	1. pull taut tissue	1 graspers	First movement of the graspers toward tissue to be cut	Termination of pulling movement; tissue taut
	2. snip tissue	1 scissors	First movement of scissors toward tissue	Closure of scissors jaws; successful separation of tissue

**Table 3.4 – Task Analysis of Laparoscopic Gallbladder Dissection
Specific Technical Skills (Joice et al 1998)**

No.	Task	Plan	No.	Subtasks
9	Detach the gallbladder from the liver bed	Do subtasks 9.1, 9.2, and 9.4 in any order then 9.4	9.1	Dissect the left medial side of the gallbladder up to fundus
			9.2	Dissect the right lateral side of the gallbladder up to the fundus
			9.3	Separate undersurface of gallbladder from liver
			9.4	Secure any bleeding from liver bed

Figure 3.4 – Hierarchy of Creation of Pneumoperitoneum (Cao et al 1999)



3.15 Conclusions

Error is part of normal human behaviour and is unavoidable in healthcare and in surgery. Surgery in particular deals with a manual task which is affected by human factors which can cause errors which can affect patient outcome.

Techniques like human reliability analysis and task analysis which have been used in commercial industries to identify and assess human errors are applicable to healthcare and surgery. In surgery as it is a complex manual task analysis is able to break down the process into quantifiable parts which can then be integrated into the whole operative process in the overall hierarchical task analysis of the operation.

This chapter has outlined the background and methods which are applicable to Surgery, and therefore in assessing the technical skills in open and laparoscopic surgery. We will use these methods as a fundamental basis in constructing and developing our assessment tools in technical skills in Surgery.

Statement of Objectives

In performing Surgery there are three main skills which are integral to the completion of an operation. Firstly, technical skills in surgery and factors which influence technical skills relating to the surgical environment. Secondly, surgical team performance and the mini-teams which work within this whole team. Finally, surgical decision making in the elective and non-elective (acute) settings.

This thesis will mainly concentrate on the technical skills assessment in non-complex simulated and live elective open and laparoscopic surgery. The operations we are concentrating on are open live primary inguinal hernia repair, live and simulated primary sapheno-femoral junction ligation and live laparoscopic cholecystectomy. We have chosen these operations because they are considered that all trainees in General Surgery should be able to perform these operations during there training period. Also they are operations all Consultants in General Surgery are expected to be competent at. We have referenced other studies and other assessment tools in technical skills to adapt and develop our own assessment tools in each of these operations. We have used two types of assessment methodology, likert and checklist scoring. For the generic technical skills assessment we have used and modified OSATS (likert) and taken and modified various core skills for the JSHST (checklist) operative competence grading. For specific technical skills we have used a combination of hierarchical task analysis and previous assessment tools to modify and develop our own likert and checklist scoring systems.

The main objectives of this research are:

1/ Technical Skills Assessment of Live Open and Laparoscopic Operations

Studies to date have concentrated mainly on bench models and virtual reality simulators to assess the technical skills of surgeons. *Assessing technical skills of surgeons on real live operations would have greater power and credibility.*

2/ Technical Skills Assessment of Expert Surgeons

The General Medical Council is presently implementing re-validation of doctors in the UK. The General Medical Council will oversee this process and re-validate expert surgeons every five years. To date there is no formal assessment of the technical skills on live open or laparoscopic surgery by expert surgeons. *So valid and reliable assessment tools in technical skills would be of great value.*

3/ Training & Self-Appraisal of Trainees & Expert Surgeons

If an assessment tool can be developed to assess generic and specific technical skills which could be modified and applied to all open and laparoscopic operations, it could be used for self-appraisal and training. *It could be used for a structured curriculum to acquire technical skills in trainees. It could also be used for expert surgeons to self-evaluate themselves.*

4/ Assessors of Technical Skills in Surgery

In previous studies there have been a range of surgeons who have assessed technical skills. *In this thesis we aim to determine the level of surgical experience required to assess technical skills of surgeons reliably and with validity.*

Methodological Study

Chapter 4

Hierarchical Task Analysis in Surgery

4.10 Overview

In this section there is an account of the methodology of how to construct a surgical task analysis and hierarchical task analysis (HTA) relating to the open and laparoscopic operations chosen. This process is the foundation required to develop an assessment tool in assessing *specific* technical skills in open and laparoscopic surgery.

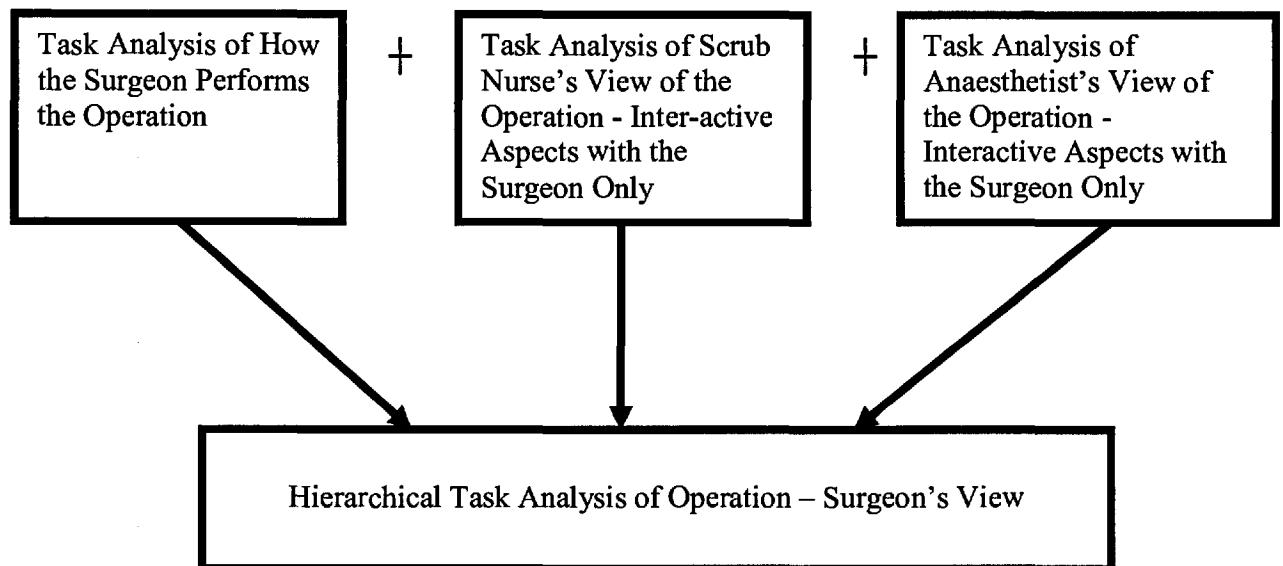
The appendix contains task analyses of the operations chosen in the view of anaesthetists and scrub nurses, which are essential in constructing the HTA of the surgical view of the operations. This is because the surgeon interacts with these two key persons when performing operations.

4.11 Rationale of Surgical Task Analysis

Surgery is a complex cascade of intricate processes working in unison. To analyse how this complex process is achieved the specific technical aspects of the operation must be evaluated in an objective, structured and quantitative way. Task analysis breaks down complex *tasks* into their component *sub-tasks* that must be performed for the task to be completed, to achieve the goals and sub-goals of the system. Sub-tasks can, of course, be further analyzed to assess by what operations they are achieved by.

Hierarchical task analysis (HTA) is a branch of task analysis where an emphasis is placed upon the organisation and members within the operating system (Kirwan ² & Ainsworth 1992). A full HTA is a combination of specific task analyses, their constitutive operations, and plans, all describing how to achieve their predecessor sub-goal, in terms of chronological completion of tasks. The hierarchy stems from the plans, which control when the operations are to be performed to complete each task. An operation cannot be done solely by a surgeon. He or she will need a surgical assistant, anaesthetist, scrub nurse etc. HTA is often represented by form of tables or flow charts, which reflect the chronological nature of this method of analysis. Diagram 4.1 outlines how a surgical HTA is constructed.

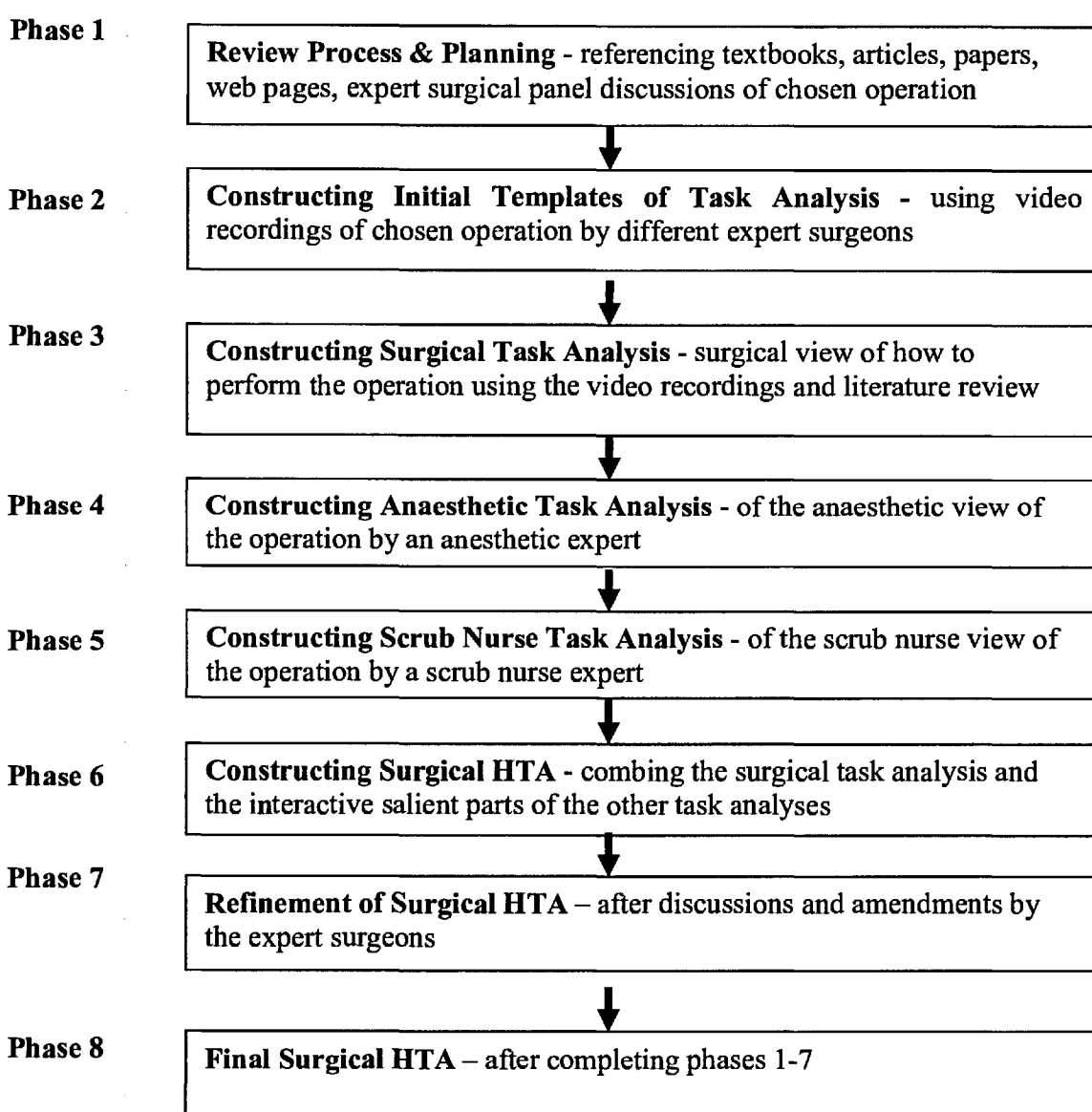
Diagram 4.1 - Constructing a Surgical Hierarchical Task Analysis for an Operation



4.12 Process of Constructing a Hierarchical Task Analysis in Surgery

To construct a surgeon's view of a HTA of an operation the following process shown in Diagram 4.2 was carried out. This HTA will be the fundamental basis of constructing an assessment tool for assessing technical skills in open and laparoscopic surgery. Each phase of the process will be described in turn.

Diagram 4.2 - Phases of Constructing a Surgical HTA for an Operation



The operations chosen are: open inguinal hernia repair, sapheno-femoral junction ligation and laparoscopic cholecystectomy, as they are performed by Trainees and Consultants in large numbers in General Surgery. They are the commonest general surgical operations in the UK and worldwide, and are fundamental operations to be learnt and performed by both Trainees and Consultants.

4.13 Phase 1 - Review Process & Planning

A review of the literature was conducted for human reliability, task analysis and HTA. All the papers written on surgical task analysis were analysed and are summarised in Chapter 3 on 'Human Error and Reliability'. After this literature review was performed, an expert surgical panel convened and the information was presented. The aim of the panel was to discuss, plan and outline the process of constructing task analysis and HTA in Surgery. The expert surgical panel discussions were then conducted every two weeks over a period of 3 months and consisted of a Consultant Surgeon, 3 Surgical Trainees and 2 experts in Human Error. After discussions it was decided that there should be a step by step process in constructing surgical task analyses and HTAs. The first step would involve recording operations by expert surgeons and using these in conjunction with papers on task analysis, to construct initial templates of task analysis of the operations.

4.14 Phase 2 – Constructing Initial Templates of Task Analysis

A prospective observational clinical study was conducted between October 2002 to March 2003 in London, UK. Patients and Consultants Surgeons were recruited at specific elective lists in the network of hospitals associated with Imperial College London. The purpose of this study was to record the essential parts of the operations performed by expert surgeons, so an initial template of the surgical task analysis could be constructed.

Ethical Approval for Video Recording

Local ethics committee approval was achieved and each patient had the opportunity for written consent and the surgeons had verbal consent. Either the patient or the surgeon had the opportunity to decline to be included in the present study.

Expert Surgeons

The expert surgeons chosen to be recorded were Consultants Surgeons who had previously performed > 150 open inguinal hernia repairs (10 expert surgeons), > 150 sapheno-femoral junction ligation (4 expert surgeons) and > 150 laparoscopic cholecystectomies (10 expert surgeons). They were a combination of General Surgeons as well as Sub-Speciality Surgeons specialising in either laparoscopic or vascular surgery. They were all deemed to be expert surgeons due to the large volume of operations they had performed successfully and independently.

Videoing

The video recordings of the operations were filmed using a Sony DCR-PC8E digital camcorder. Full length versions of operations were recorded onto digital video tape. These were then transferred without editing onto DVD, using Sony Click to DVD Software (Tokyo, Japan).

Open Inguinal Hernia Group of Expert Surgeons

20 video recordings of primary open inguinal hernia repairs from 10 different Consultants Surgeons (2 cases per Consultant Surgeon) were assessed in view of constructing a task analysis.

Sapheno-Femoral Junction Ligation Group of Expert Surgeons

8 video recordings of sapheno-femoral junction ligation from 4 different Consultants Surgeons (2 cases per Consultant Surgeon) were assessed in view of constructing a task analysis.

Laparoscopic Cholecystectomy Group of Expert Surgeons

20 video recordings of laparoscopic cholecystectomy from 10 different Consultants Surgeons (2 cases per Consultant Surgeon) were assessed in view of constructing a task analysis.

Initial Template of Surgical Task Analysis

Defining Fundamental Aspects of the Operations

Two surgeons (the researchers) with >12 years postgraduate surgical experience (Consultant Surgeon and Specialist Registrar) assessed each operation blindly and independently. The operations were watched on large screen televisions. All of the operations were seen. Fundamental tasks of the operations were noted by the researchers in consecutive order. Several videos from different experts were viewed as each expert surgeon has his or her individual technical style. Table 4.1 demonstrates the fundamental tasks of successfully performing for example, a laparoscopic cholecystectomy with the consecutive order of the steps. From this, initial templates of task analyses of the operations were constructed.

Table 4.1 – Initial Template of Task Analysis of a Laparoscopic Cholecystectomy

Fundamental Tasks	Order of Step
Creation of Pneumo-peritoneum	1
Ports Insertion	2
Gall bladder retraction	3
Exposure & Dissection of Calot's triangle	4
Cystic Artery Dissection	5
Cystic Duct Dissection	6
Cystic Artery Clipping & Transection	7
Cystic Duct Clipping & Transection	8
Gall bladder Fossa Dissection	9
Extraction of gallbladder	10
Closure of Ports	11

4.15 Phase 3 – Constructing Surgical Task Analysis

Procedure

A standard task analysis of the operations were constructed by using video recordings of the expert surgeons and the initial task analysis templates. Each fundamental task (step) described in Table 4.1 for example, was then divided into sub-tasks and recovery steps. Table 4.2 demonstrates an example of a task in performing a laparoscopic cholecystectomy with its sub-tasks and possible recovery steps which leads to the successful completion of the task. The recovery aspects are required as tasks in Surgery are dynamic and sub-tasks may be repeated or altered to achieve the goal.

Table 4.2 – Task Analysis of One Task in a Laparoscopic Cholecystectomy

No.	Task	Plan	No.	Sub-Tasks	Recovery
5	Laparoscopy, retraction, dissect & expose Calot's triangle (cystic artery, cystic duct)	Do subtasks 5.1, 5.2, 5.3 in consecutive order	5.1	Laparoscopy of abdomen	Abnormal intra-abdominal pathology, decide to continue or not.
			5.2	Retract gall bladder	
			5.3	Dissect adhesions to gallbladder	Bleeding from port site – rectify before proceeding
			5.4	Dissect & mobilise Hartmann's pouch	Graspers detach from gallbladder, reapply graspers
			5.5	Dissect & isolate cystic duct	
			5.6	Dissect & mobilise cystic artery	Any significant bleeding during subtasks 5.3 to 5.6 will need correcting

4.16 Phase 4 - Constructing an Anaesthetic Task Analysis

When performing an operation the surgeon does not perform this without any help. In performing an operation under general anaesthesia, the surgeon interacts with an anaesthetist. Therefore it is important to define a task analysis of the anaesthetist of the operation, and identify when the surgeon interacts specifically with the anaesthetist. This

will enable us to construct a HTA for the operation. A senior Consultant Anaesthetist constructed a task analysis for each of the chosen operations from an anaesthetist's point of view (Appendix - Tables 4.13, 4.15, 4.17). The anaesthetist chosen had been a Consultant Anaesthetist for more than 10 years. The expert anaesthetist was given instructions and advice how to construct a task analysis. Table 4.3 demonstrates a task where the surgeon and anaesthetist interact in performing a laparoscopic cholecystectomy.

Table 4.3 – Task Analysis of One Task Demonstrating Surgeon & Anaesthetist Interaction in a Laparoscopic Cholecystectomy

No.	Task	Plan	No.	Sub-Tasks	Recovery
6	Rectify anaesthetic or operative difficulty (not all cases and can be at any part of operation)	Do subtasks 6.1, 6.2, 6.3 if required	6.1	Speak to anaesthetist	Wait until anaesthetist gives approval to continue if anaesthetic problem
			6.2	Halt operation if necessary or rectify operative problem	
			6.3	Continue with operation	Major operative problem, conversion to open cholecystectomy 6.1 to 6.2 may need several repetitions before 6.3 can be achieved

4.17 Phase 5 – Constructing a Scrub Nurse Task Analysis

When performing an operation the surgeon does not perform this without any help. In performing an operation the surgeon interacts with the scrub nurse. Therefore it is important to define a task analysis of the scrub nurse of the operations, and identify when the surgeon interacts specifically with the scrub nurse. This will enable us to construct a HTA for the operation. A senior scrub nurse constructed a task analysis for each of the chosen operation from a scrub nurse's view. (Appendix - Tables 4.14, 4.16, 4.18). The scrub nurse chosen had been a senior nurse for more than 10 years. The expert scrub

nurse was given instructions and advice how to construct a task analysis. Table 4.4 demonstrates a task where the surgeon and scrub nurse interact in performing a laparoscopic cholecystectomy.

Table 4.4 – Task Analysis One Task Demonstrating Surgeon & Scrub Nurse Interaction in a Laparoscopic Cholecystectomy

No.	Task	Plan	No.	Sub-Tasks	Recovery
16	Check with scrub nurse swabs & instruments are correct	Do subtasks 16.1, 16.2	16.1	Ask scrub nurse that swab & instrument count is correct	Swab & instrument count incorrect.
			16.2	Acknowledge to scrub nurse that you have heard her or him	Find missing item(s) and reevaluate count.

4.18 Phase 6 – Constructing Surgical Hierarchical Task Analysis

Surgical HTAs of each operation were then constructed using the task analyses of the surgical view by the expert surgeons of the chosen operations, and combining with those parts which complement the surgeon from the task analyses of the anaesthetist and scrub nurse (Appendix - Tables 4.10, 4.11, 4.12).

4.19 Phase 7 – Refining Surgical Hierarchical Task Analysis

Procedure

The aim of this present phase was to refine the HTA constructed in phase 6, and to construct a consensus HTA of each operation (tasks and sub-tasks). The eventual HTA should be descriptive and not prescriptive, so any surgeon could use it. So these preliminary HTA of the operations, constructed in phase 6, were then given to each group of Consultant Surgeons performing these operations. The HTA were evaluated and modified if required by each of the Consultant Surgeons participating to assess whether it

differed from their own prescribed set of tasks for completion of the operation. The HTA was performed to a level which described the tasks and sub-tasks which had to be completed to achieve the goal, but did not describe the technique and instruments that should be used. This was done so that the individual technical style of the surgeons was unbiased by any implied technique. Reliability analysis was done between each individual Consultant Surgeon and the Researchers, who constructed the initial HTA in phase 6. This was done to demonstrate if there was a reliable consensus to the final construction of the HTA.

Statistical Analysis – Reliability Between Researchers & Expert Surgeons

Data was collated in an Excel[®] database (Microsoft, Redmond, Washington, USA). Statistical analysis was carried out with SPSS[®] software statistical package (SPSS, Chicago, Illinois, USA). The HTA in phase 6 was compared to each individual Consultant Surgeons HTA. The comparison was done using a checklist method, i.e. each task and sub-task was marked either the same, similar or different than that in phase 6, compared to each individual surgeon (Table 4.5). A total score was then generated for each individual surgeon (the researchers having the maximum score); the score was then converted to a percentage score compared to the researchers.

Table 4.5 – Checklist Scoring for Percentage Agreement Between Researchers & Individual Expert Surgeons – Tasks & Sub-Tasks

Score	Level of Agreement
4	Yes – Complete (100%) agreement between researchers and individual expert surgeon
2	Yes – Some (> 50%) agreement between researchers and individual expert surgeon
0	No – No (< 50%) agreement between researchers and individual expert surgeon

Results

Open Inguinal Hernia Group of Expert Surgeons

Each Consultant Surgeon assessed the HTA and modified it according to their own technical style. For *tasks* there was a 100% correlation between the researchers and individual expert surgeons. There were 47 sub-tasks with a total possible score of 188 (researchers score) using the scoring system in Table 4.5. The total scores for each surgeon are shown in Table 4.1A. The scores were converted to percentages of agreement. Mean percentage agreement for *sub-tasks* was 84%, (range 73-90%). Figure 4.1 demonstrates the percentage agreement of the 10 Consultant Surgeons compared to the original HTA *sub-tasks* of the researchers in phase 6. The final HTA which was used to develop the tool for assessing technical skills used the amalgamated HTAs (tasks & sub-tasks) of all the Consultant Surgeons.

Figure 4.1 Percentage Agreement Between Expert Surgeons and Researchers of Sub-Tasks of Primary Inguinal Hernia Repair

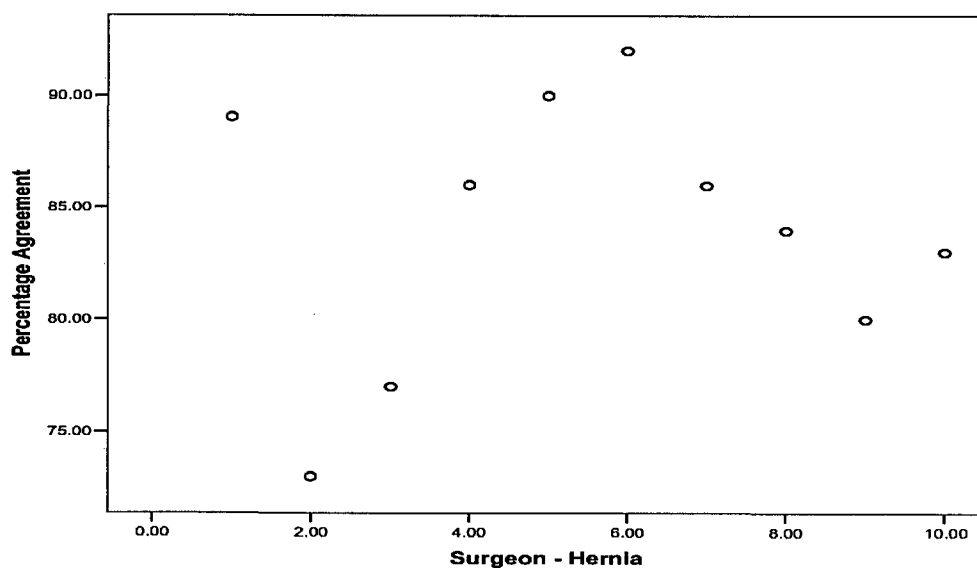


Table 4.1A - Total Scores for each Surgeon for Sub-Tasks – Inguinal Hernia

Surgeon	1	2	3	4	5	6	7	8	9	10
Score	165	136	145	161	169	174	161	160	147	155

Sapheno-Femoral Junction Ligation Group of Expert Surgeons

Each Consultant Surgeon assessed the HTA constructed by the researchers and modified it according to their own technical style. For *tasks* there was a 100% correlation between the researchers and individual expert surgeons. There were 52 sub-tasks with a total possible score of 208 (researchers score) using the scoring system in Table 4.5. The total scores for each surgeon are shown in Table 4.2A. The scores were converted to percentage agreement. Mean percentage agreement for *sub-tasks* was 81% (range 75-90%). Figure 4.2 demonstrates the percentage agreement of the 4 Consultant Surgeons compared to the original HTA *sub-tasks* of the researchers in phase 6. The final HTA

which was used to develop the tool for assessing technical skills used the amalgamated HTAs (tasks & sub-tasks) of all the Consultants.

Figure 4.2 Percentage Agreement Between Expert Surgeons and Researchers of Sub-Tasks of Primary Sapheno-Femoral Junction Ligation

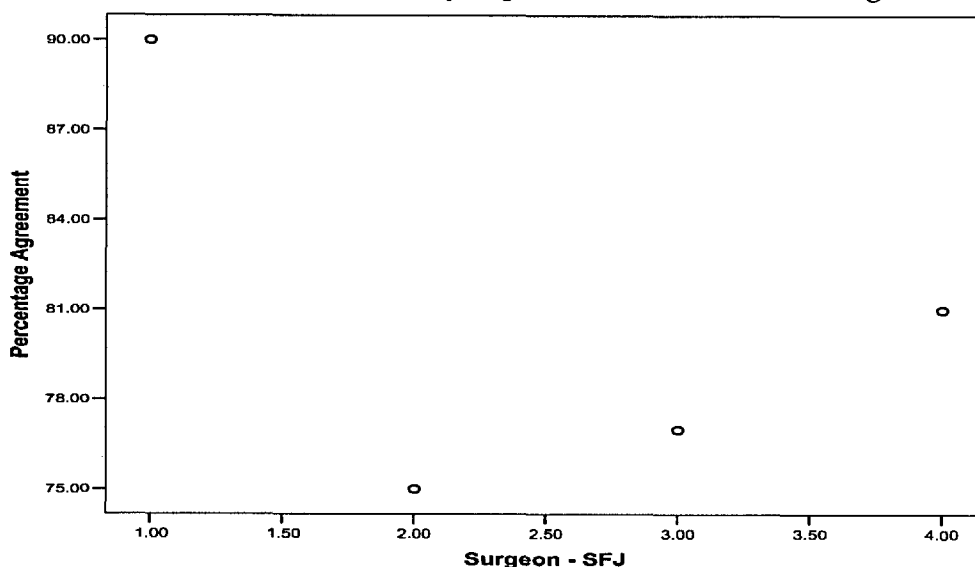


Table 4.2A - Total Scores for each Surgeon for Sub-Tasks – Sapheno-Femoral Junction

Surgeon	1	2	3	4
Score	187	156	160	168

Laparoscopic Cholecystectomy Group of Expert Surgeons

Each Consultant Surgeon assessed the HTA constructed by the researchers and modified it according to their own technical style. For *tasks* there was a 100% correlation between the researchers and individual expert surgeons. There were 53 sub-tasks with a total possible score of 212 (researchers score) using the scoring system in Table 4.5. The total scores for each surgeon are shown in Table 4.3A. The scores were converted to percentage agreement. Mean percentage agreement for *sub-tasks* was 84% (range 73-92%). Figure 4.3 demonstrates the percentage agreement of the 10 Consultant Surgeons

compared to the original HTA *sub-tasks* of the researchers in phase 6. The final HTA which was used to develop the tool for assessing technical skills used the amalgamated HTAs (tasks & sub-tasks) of all the Consultants.

Figure 4.3 Percentage Agreement Between Expert Surgeons and Researchers of Sub-Tasks of Laparoscopic Cholecystectomy

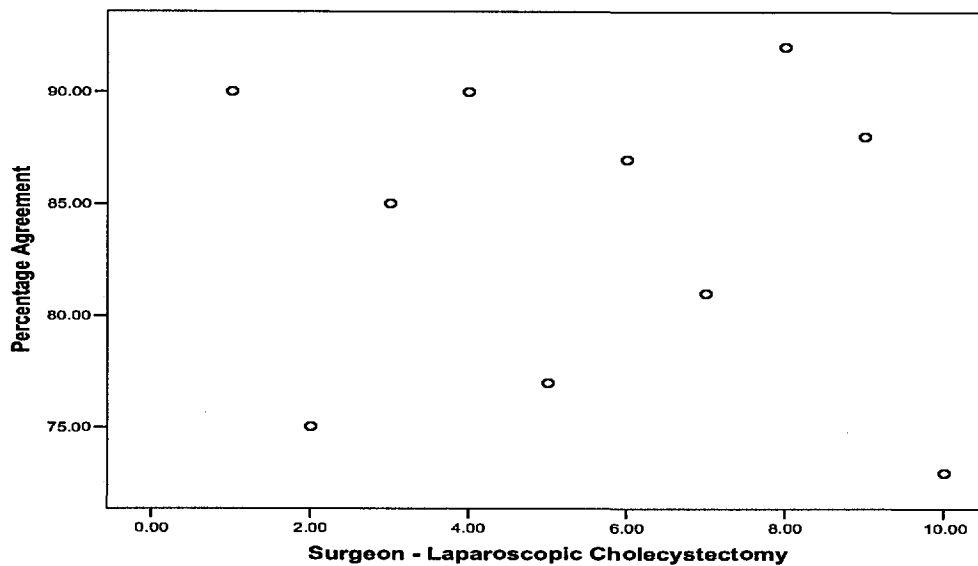


Table 4.3A - Total Scores for each Surgeon for Sub-Tasks – Laparoscopic Cholecystectomy

Surgeon	1	2	3	4	5	6	7	8	9	10
Score	191	159	180	191	164	186	171	195	187	154

4.20 Phase 8 – Constructing Final Surgical Hierarchical Task Analysis

Final Surgical HTA of each operation were then constructed using the amalgamated HTAs in phase 7. This was mainly incorporating the different technical styles of each individual expert surgeon which was reflected in the sub-tasks, which had some variability. The final HTAs of each operation are shown in Tables 4.10, 4.11, 4.12.

Table 4.10 HTA of Primary Inguinal Hernia Repair – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
1	Position patient, prep & drape inguinal area	Do subtasks 1.1,1.2 in consecutive order	1.1 1.2	Position patient supine Prep inguinal area Drape inguinal area	Repeat 1.1 to 1.2 if necessary in consecutive order if any subtasks fails
2	Inform the anaesthetist that you (the surgeon) intends to start operation	Do subtasks 2.1 to 2.2	2.1 2.2	Speak to anaesthetist that you intend to start Acknowledge anaesthetist has given approval	Anaesthetist says not to start, wait for approval and repeat 2.1 to 2.2 in consecutive order
3	Incision in inguinal area	Do subtasks 3.1,3.2 in consecutive order	3.1 3.2	Check anatomical surface landmarks (anterior superior iliac spine and pubic tubercle i.e. inguinal ligament) Transverse incision 3cm above inguinal ligament	Bleeding from superficial veins Stop with diathermy
4	Identify external oblique aponeurosis & external ring	Do subtasks 4.1,4.2 in consecutive order	4.1 4.2	Dissect down onto external oblique aponeurosis Identify external ring	Initial incision not exact, dissect to compensate and achieve subtasks
5	Open up inguinal canal	Do subtasks 5.1, 5.2, 5.3, 5.4 in consecutive order	5.1 5.2 5.3 5.4	Small incision transversely on external oblique aponeurosis in line with external ring Incision extended laterally and medially releasing external ring Superior & inferior leaves dissected off & clips applied to each leaf, lateral pocket formed Travers retractor inserted in between superior & inferior leaves	Travers retractor comes off, reapply retractor
6	Rectify anaesthetic or operative difficulty (not all cases and can be at any part of operation)	Do subtasks 6.1, 6.2, 6.3 if required	6.1 6.2 6.3	Speak to anaesthetist Halt operation if necessary or rectify operative problem Continue with operation	Wait until anaesthetist gives approval to continue if anaesthetic problem Rectify operative problem
7	Identify spermatic cord	Do subtasks 7.1, 7.2 in consecutive order	7.1 7.2	Sweep fingers around spermatic cord medially onto pubic tubercle Dissect fibres off posterior aspect of spermatic cord	Spermatic cord severely damaged may proceed to orchidectomy
8	Identify indirect or direct sac of inguinal hernia	Do task 8		If direct hernia reduce hernia & plicate posterior wall (transversalis fascia) If indirect hernia dissect hernial sac from spermatic cord superior-medially, preserving cord contents	Bleeding from cord or surrounding area, stop with diathermy preserving cord contents.

Table 4.10 HTA of Primary Inguinal Hernia Repair – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
9	Removal of indirect inguinal hernia sac	Do subtasks 9.1, 9.2, 9.3, 9.4, 9.5 in consecutive order	9.1 9.2 9.3 9.4 9.5	Open indirect hernia sac Reduce contents Suture ligate neck of sac Excise sac Reduce neck of sac back into abdominal cavity	Viscera injured whilst opening sac. Repair damage and proceed as appropriate
10	Placement of mesh	Do subtasks 10.1,10.2,10.3,10.4,10.5, 10.6, 10.7 in consecutive order	10.1 10.2 10.3 10.4 10.5 10.6 10.7	Cut mesh to shape Retract spermatic cord medially Place stitch medial to pubic tubercle Place mesh under spermatic cord Stitch mesh along inguinal ligament Interrupted stitches medially, superiorly and around new mesh internal ring Place lateral aspect of mesh into lateral pocket	Mesh not cut appropriately or became non-sterile. Request for new mesh and cut to appropriate shape.
11	Inform anaesthetist that the mesh is in place and the operation is drawing to an end	Do subtasks 11.1, 11.2	11.1 11.2	Speak to anaesthetist Acknowledge that anaesthetist has heard you	
12	Removal of retractors	Do subtasks 12.1, 12.2 in consecutive order	12.1 12.2	Remove retraction around spermatic cord Remove Travers retractor around superior & inferior leaves of external oblique aponeurosis	
13	Closing external oblique aponeurosis	Do subtasks 13.1, 13.2, 13.3, 13.4, 13.5, 13.6 in consecutive order	13.1 13.2 13.3 13.4 13.5 13.6	Place Travers retractor around skin edges Lift clips anteriorly Stitch superior& inferior leaves together Remove clips from superior & inferior leaves half-way during stitching Complete stitching laterally Remove Travers retractor	Bleeding from subcutaneous tissues Stop bleeding with diathermy
14	Injection of local anaesthetic	Do subtasks 14.1, 14.2 in consecutive order	14.1 14.2	Check local anaesthetic agent Inject local anaesthetic agent around ilio-inguinal nerve	Wrong local anaesthetic, replace with correct one

Table 4.10 HTA of Primary Inguinal Hernia Repair – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
15	Check with scrub nurse swabs & instruments are correct	Do subtasks 15.1, 15.2	15.1 15.2	Ask scrub nurse that swab & instrument count is correct Acknowledge to scrub nurse that you have heard her or him	Swab & instrument count incorrect. Find missing item(s) and reevaluate count.
16	Close up patient	Do subtasks 16.1, 16.2, 16.3 in consecutive order	16.1 16.2 16.3	Suture skin wound Clean wound Place dressing over wound	Dressings fall off due to wet skin surrounding wound. Dry surrounding skin and reapply dressing.

Table 4.11 HTA of Sapheno-Femoral Junction Ligation – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
1	Position patient, prep & drape femoral area	Do subtasks 1.1, 1.2, 1.3 in consecutive order	1.1 1.2 1.3	Position patient supine, with leg externally rotated and abducted on a vein board Prep femoral area Drape femoral area	Repeat 1.1 to 1.3 if necessary in consecutive order if any subtasks fails
2	Inform the anaesthetist that you (the surgeon) intends to start operation	Do subtasks 2.1 to 2.2	2.1 2.2	Ask anaesthetist if you can start Acknowledge anaesthetist has given approval	Anaesthetist says not to start, wait for approval and repeat 2.1 to 2.2 in consecutive order
3	Incision in femoral area	Do subtasks 3.1, 3.2 in consecutive order	3.1 3.2	Check anatomical surface landmarks (below pubic tubercle & inguinal ligament in groin crease) Transverse incision in groin	Bleeding from superficial veins Stop with diathermy or ties
4	Dissect down Scarpa's & cribriform fascia	Do tasks 4.1, 4.2, 4.3 in consecutive order	4.1 4.2 4.3	Dissect down Scarpa's & cribriform fascia Identify long saphenous vein emerging from femoral vein Travers retractor inserted in wound to expose sapheno-femoral junction area	Initial incision not exact, dissect to compensate and achieve subtasks Bleeding from superficial veins. Stop with diathermy or ties Travers retractor comes off, reapply retractor
5	Identifying & tying off all known tributaries (4) (Superficial circumflex iliac, Superficial inferior epigastric, Superficial & Deep external pudal) & any others	Do subtasks 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7 in consecutive order	5.1 5.2 5.3 5.4 5.5 5.6 5.7	Identify tributary, dissect around for clearance Pass haemostat under tributary and grip tie & pass under tributary Knot tie proximally onto tributary Pass haemostat under tributary and grip tie & pass under tributary Knot tie distally onto tributary Cut between ties Repeat subtasks 5.1 to 5.6 until all tributaries are tied	After cutting tributary, bleeding from either cut end. Apply haemostat clip to distal end and re-knot distal end with new tie

Table 4.11 HTA of Sapheno-Femoral Junction Ligation – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
6	Rectify anaesthetic or operative difficulty (not all cases and can be at any part of operation)	Do subtasks 6.1, 6.2, 6.3 if required	6.1 6.2 6.3	Speak to anaesthetist Halt operation if necessary or rectify operative problem Continue with operation	Wait until anaesthetist gives approval to continue if anaesthetic problem Rectify problem
7	Identify & ligate sapheno-femoral junction	Do subtasks 7.1, 7.2, 7.3, 7.4, 7.5, 7.6 in consecutive order	7.1 7.2 7.3 7.4 7.5 7.6	Dissect around sapheno-femoral junction. Identify & sapheno-femoral junction Suture ligate sapheno-femoral junction not encroaching onto femoral vein Knot tie around long saphenous vein distally from sapheno-femoral junction tie Place clip on proximal part of long saphenous vein Cut between two ties	Bleeding from ligated end – reclamp and retie Femoral vein damaged. Request for Yanker sucker and vascular clamps. Repair damage with appropriate suture. If still uncontrollable bleeding call for senior or vascular help
8	Stripping of long saphenous vein	Do subtasks 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 8.10, 8.11, 8.12 in consecutive order	8.1 8.2 8.3 8.4 8.5 8.6 8.7 8.8 8.9 8.10 8.11 8.12	Cut a small hole in proximal part of long saphenous vein Pass stripper down long saphenous to just below knee Make transverse cut in skin & long saphenous vein over distal part of stripper Apply clip to distal part of stripper and long saphenous vein Apply clip to distal part of exposed long saphenous vein Transect long saphenous vein between clips Knot tie distal clip Re-apply clip to stripper and pull it out from wound Place appropriate size head to distal part of stripper Pull stripper from knee towards groin, stripping long saphenous vein Apply pressure to tract of long saphenous vein in thigh Milk excess blood & haematoma from tract through groin wound	Clip falls off. Reapply clip. Tie falls off. Reapply tie. Distal skin wound too small, enlarge so head of stripper can pass through.

Table 4.11 HTA of Sapheno-Femoral Junction Ligation – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
9	Multiple phlebectomies (Not done in bench model)	Do subtasks 9.1, 9.2, 9.3, 9.4,9.5 in consecutive order	9.1 9.2 9.3 9.4 9.5	Ask patient to be put in Trendelenberg position Make longitudinal cuts over marked varicose veins Feed hook or clip around varicose vein Tease varicose vein out reapplying clips as necessary Apply pressure with a small swab over phlebectomy site	Lumen of varicose vein breached causing bleeding. Apply pressure over site.
10	Inform anaesthetist that most of the phlebectomies are done and the operation is drawing to an end	Do subtasks 10.1, 10.2	10.1 10.2	Speak to anaesthetist Acknowledge that anaesthetist has heard you	
11	Check with scrub nurse swabs & instruments are correct	Do subtasks 11.1, 11.2	11.1 11.2	Ask scrub nurse that swab & instrument count is correct Acknowledge to scrub nurse that you have heard her or him	Swab & instrument count incorrect. Find missing item(s) and reevaluate count.
12	Close up patient	Do subtasks 12.1, 12.2, 12.3, 12.4, 12.5 in consecutive order	12.1 12.2 12.3 12.4 12.5	Suture groin and distal skin wounds Clean wounds Elevate leg Place dressings over wounds Apply crepe bandage firmly around leg up to thigh	Dressings fall off due to wet skin surrounding wound. Dry surrounding skin and reapply dressings.

Table 4.12 HTA of Laparoscopic Cholecystectomy – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
1	Position patient, prep & drape abdomen, position & attach laparoscopic instruments and equipment	Do subtasks 1.1,1.2,1.3,1.4 in consecutive order	1.1 1.2 1.3 1.4	Position patient supine Prep abdomen Drape abdomen Position & attach laparoscopic instruments and equipment (camera, gas and diathermy leads, monitor & camera)	Repeat 1.1 to 1.4 if necessary in consecutive order if any subtasks fails
2	Inform the anaesthetist that you (the surgeon) intends to start operation	Do subtasks 2.1 to 2.2	2.1 2.2	Speak to anaesthetist that you intend to start Acknowledge anaesthetist has given approval	Anaesthetist says not to start, wait for approval and repeat 2.1 to 2.2 in consecutive order
3	Create CO2 pneumoperitoneum	Do subtasks 3.1,3.2,3.3 in consecutive order	3.1 3.2 3.3	Open technique Insert 10mm umbilical port without trocar Insufflate abdomen with CO2	Pneumoperitoneum not created repeat 3.1 to 3.3 in consecutive order
4	Insert laparoscopic ports	Do subtasks 4.1,4.2,4.3 in consecutive order	4.1 4.2 4.3	Insert 10mm port (epigastric) Insert 5mm port (lateral) Insert 5mm port (lateral)	Ports not inserted correctly. Repeat 4.1, 4.2, 4.3 as required
5	Laparoscopy, retraction, dissect & expose Calot's triangle (cystic artery, cystic duct)	Do subtasks 5.1, 5.2, 5.3 in consecutive order	5.1 5.2 5.3 5.4 5.5 5.6	Laparoscopy of abdomen Retract gall bladder Dissect adhesions to gallbladder Dissect & mobilise Hartmann's pouch Dissect & isolate cystic duct Dissect & mobilise cystic artery	Abnormal intra-abdominal pathology, decide to continue or not. Bleeding from port site – rectify before proceeding Graspers detach from gallbladder, reapply graspers Any significant bleeding during subtasks 5.3 to 5.6 will need correcting
6	Rectify anaesthetic or operative difficulty (not all cases and can be at any part of operation)	Do subtasks 6.1, 6.2, 6.3 if required	6.1 6.2 6.3	Speak to anaesthetist Halt operation if necessary or rectify operative problem Continue with operation	Wait until anaesthetist gives approval to continue if anaesthetic problem Major operative problem, conversion to open cholecystectomy 6.1 to 6.2 may need several repetitions before 6.3 can be achieved
7	Secure cystic artery & cystic duct	Do subtasks 7.1, 7.2, 7.3 in any order	7.1 7.2 7.3	Place 2 clips on proximal end of cystic artery Place clip on distal end of cystic artery Place clip at gallbladder end of cystic duct	Clips are not placed correctly or fall into abdomen. Retrieve & remove incorrect clips. Reapply clips 7.1 to 7.3 as necessary

Table 4.12 HTA of Laparoscopic Cholecystectomy – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
8	Divide cystic artery between clips	Do task 8			
9	Perform operative cholangiogram (not all cases)	Do subtasks 9.1, 9.2, 9.3, 9.4, 9.5 in consecutive order	9.1 9.2 9.3 9.4 9.5	Incise antero-superior wall of cystic duct Insert cholangiogram catheter into cystic duct Inject contrast – fluorocholangiogram Shoot cholangiogram Remove catheter	Catheter falls out, repeat 9.2 Cholangiogram fails or of poor quality, repeat 9.3 & 9.4 in consecutive order
10	Secure proximal end of cystic duct	Do 10.1	10.1	Place 2 clips on proximal end of cystic duct and proximal to incision on cystic duct	Clips are not placed correctly or fall into abdomen. Retrieve & remove incorrect clips. Reapply clips.
11	Divide cystic duct between clips	Do task 11			
12	Dissect gallbladder from liver bed	Do subtasks 12.1, 12.2, & 12.3 in any order then 12.4	12.1 12.2 12.3 12.4	Dissect left (medial) side of gallbladder up to fundus Dissect right (lateral) side of gallbladder up to fundus Separate under surface of gall bladder from liver Secure any bleeding from liver bed	Excessive bleeding, irrigate area and diathermy where necessary. Placement of Surgicel™ or similar material
13	Removal of dissected gallbladder	Do subtasks 13.1, 13.2, 13.3 in consecutive order	13.1 13.2 13.3	Insert retrieval bag Place gallbladder inside bag Extract bag containing gallbladder	Bag bursts, remove bag & repeat 13.1 to 13.3 in consecutive order. 10mm port too small for gallbladder and bag. Enlarge port site or remove gallstones from gallbladder before doing 13.3
14	Inform anaesthetist the gallbladder is out and the operation is drawing to an end	Do subtasks 14.1, 14.2	14.1 14.2	Speak to anaesthetist Acknowledge that anaesthetist has heard you	

Table 4.12 HTA of Laparoscopic Cholecystectomy – Surgical View

No.	Task	Plan	No.	Sub-Tasks	Recovery
15	Final check (all cases) irrigation & placement of drain (not all cases)	Do subtasks 15.1, 15.2, 15.3, 15.4, 15.5 in any order	15.1 15.2 15.3 15.4 15.5	Check & coagulate any bleeding areas Check cystic artery stump & clips Check cystic duct & clips Irrigate & suction operative field Place drain under liver bed	Drain not placed correctly. Repeat 15.5
16	Check with scrub nurse swabs & instruments are correct	Do subtasks 16.1, 16.2	16.1 16.2	Ask scrub nurse that swab & instrument count is correct Acknowledge to scrub nurse that you have heard her or him	Swab & instrument count incorrect. Find missing item(s) and reevaluate count.
17	Close up patient	Do subtasks 17.1, 17.2, 17.3, 17.4, 17.5 in consecutive order 17.6, 17.7	17.1 17.2 17.3 17.4 17.5 17.6 17.7	Remove epigastric and lateral ports Check port sites Release CO2 from abdomen Remove umbilical port Suture port sites Clean port site areas Place dressings over port sites	Bleeding from port site. Rectify before proceeding with subtasks 17.3 to 17.7

4.21 Conclusions

A surgical HTA of an operation demonstrates how the task analyses of a surgeon, anaesthetist and scrub nurse interacts with each other in proceeding, progressing and finally completing the operation, which is the main goal. In this present study we have concentrated on constructing surgical HTAs on key operations performed by General Surgeons. We therefore constructed surgical HTAs for the following operations: primary inguinal hernia repair, sapheno-femoral junction ligation, laparoscopic cholecystectomy.

The percentage agreement between the expert surgeons' HTA and the researchers' HTA for all three operations were good. By constructing surgical HTAs, an operation can be systematically broken down into individual components (procedural tasks and sub-tasks) which can then be potentially quantitatively assessed. These key components of the surgical HTAs will now be analysed and aspects of these will be used to develop an assessment tool to assess the specific technical skills in these specified open and laparoscopic operations.

Constructing an assessment tool for assessing specific technical skills in open and laparoscopic surgery must be based on a consensus of expert opinions and must have good agreement. In this present study we have demonstrated both. The template of the tool to assess specific technical skills will be broad and will assess procedure specific and minor and major technical errors relating to the individual operations. The next chapter will outline how the assessment tool for specific technical skills was developed from the surgical HTAs.

Methodology

Chapter 5

Developing Criteria to Assess Generic & Specific Technical Skills in Surgery

5.10 Overview

In the previous chapter on hierarchical task analysis we were able to break down the chosen operations into a step by step process. In this chapter we will use the hierarchical task analyses to develop likert and checklist scores for specific technical skills. For the likert and checklist generic technical skills we will adapt existing scoring systems (OSATS & JCHST). When developing an assessment scoring system various aspects of the scoring system need to be clarified. Finally various aspects of likert and checklist scoring will be discussed which show their strengths and weaknesses.

5.11 Methods - Generic Technical Skills

Likert Scoring

The generic technical skills likert scoring OSATS (Objective Structured Assessment of Technical Skills) was developed by a group in Toronto (Martin ¹ et al 1995). It has been extensively been validated in numerous studies over the last 10 years (Chapter 2).

This type of system is subjective as it has a range of possible scores. However it is more sensitive and has the ability to discriminate varying abilities within the same group as

well as with different groups. As this type of scoring is a continuum, parametric analysis can be used in statistical analysis (Bland J.M. 2000, Armitage P. et al 2002).

However it only assesses the generic technical skills used by the surgeon in performing the operation, and not the quality of the specific steps of the procedure or operation. We adapted and modified the version of OSATS (Table 5.1A). deleting sections e.g. 'knowledge of the procedure', as this was more comprehensively covered in the operation specific technical skills scale Other categories had their wording changed e.g. 'use of assistants' to 'exposure of technical field'. Generic likert technical skills were scored incrementally by 1 point. The new likert scale is shown in Table 5.1B

Table 5.1A – Generic Technical Skills – OSATS Likert Scoring Assessment Tool

Category	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 & 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 the names of most instruments and used appropriate instrument or the task		Obviously familiar with the instruments required and their names
Use of Assistants	Consistently placed assistants poorly or failed to use assistants		Good use of assistants most of the time		Strategically used assistant to the best advantage at all times
Flow of Operation & 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 aspects of the operation		Demonstrated familiarity with all aspects of the operation

Table 5.1B – Generic Technical Skills – PhD Likert Scoring Assessment Tool

Category	1	2	3	4	5
Handling & Dissection of Tissues	Unnecessary force on tissues, causing frequent damage to tissues		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time & Motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument & Material Handling	Repeatedly makes tentative or awkward moves with instruments or material		Competent use of instruments/material although occasionally awkward		Fluid moves with instruments/material and no awkwardness
Use of Instruments & Materials	Frequently used inappropriate instruments & materials		Used appropriate instrument & materials for the task or subtask		Expertly used appropriate instruments & materials
Exposure of Technical Field	Consistently poor exposure		Good exposure most of the time		Expert exposure
Completion & Flow of Tasks & Subtasks	Frequently stopped operating or hesitant in next task or subtask		Forward planning with steady progression of tasks & subtasks		Effortless flow from one task to subtask

Checklist Scoring

This type of scoring system is finite and has definite yes/no type of scoring. The benefit of this scoring system in Surgery is that it can assess if technical errors have been committed, as these are specific events during the operation. As this type of scoring is finite, non-parametric analysis is used in statistical analysis.

An assessment system of core generic skills was developed by the Joint Committee in Higher Surgical Training (www.jchst.org) in the UK (Table 5.2A) as a guideline to use the Operative Competence Grading System (Table 2.5). Although it has not been validated, it has been recommended for assessing surgical technical skills in higher surgical training

We have adapted and changed this checklist of core skills in this thesis to a checklist of generic technical skills scoring system, deleting and modifying as necessary (Table 5.2B). Generic technical skills errors were scored 0 point if correctly done and 1 point if incorrectly done. Aspects of this checklist mirror the generic technical skills in our new likert scoring system, and is useful to determine whether likert or checklist scoring is better at assessing generic technical skills.

Table 5.2A – Generic Core Skills – JCHST Assessment Tool

Category		
	Yes	No
Checks patient case record and that consent has been obtained		
Communicates well with the theatre team		
Thorough preparation (marking, catheter, antibiotics etc)		
Good scrub and aseptic technique		
Positions patient correctly on operating table		
Makes appropriate incision (s)		
Purposeful dissection in correct tissue planes		
Demonstrates sound knowledge of anatomy		
Familiar with all steps of the procedure		
Uses the correct instruments efficiently		
Handles dangerous instruments safely		
Uses assistant (s) to best advantage		
Possesses good hand-eye co-ordination		
Handles tissues gently and dextrously		
Reliable suturing and knotting techniques		
Sound repair or anastomosis		
Uses diathermy appropriately and safely		
Able to control bleeding by suction, clips or sutures		
Closes wound neatly and securely		
Timeliness: procedure unhurried with no unnecessary delay		
Good documentation (operation note and postoperative instructions)		

Table 5.2B – Generic Technical Skills – PhD Checklist Scoring Assessment Tool

Category	0	1
	Yes	No
Appropriate access to technical field		
Correct understanding of tissue planes		
Demonstrates sound knowledge of anatomy		
Confident & efficient progression of tasks & subtasks		
Uses instruments & materials efficiently & correctly		
Appropriate exposure/retraction of technical field		
Possesses good hand-eye co-ordination		
Handles tissues gently with minimal damage		
Correct suturing and knotting techniques		
Uses diathermy appropriately and safely		
Able to control bleeding efficiently & competently		
Appropriate wound management - closure & dressings		
Tasks or subtasks not rushed or too slow		

5.12 Methods - Specific Technical Skills

Likert Scale - Assessing Key Tasks in Surgery

The likert scoring system was developed by identifying the key steps or tasks in the operation by using the hierarchical task analyses developed in Chapter 4. Below we have demonstrated how this was done for laparoscopic cholecystectomy as an example. Tables 5.3, 5.4, 5.5, 5.6, 5.7 show the five key steps in this operation, and from this we developed the likert scoring system (Table 5.8) after the refinement and feasibility study was done (see below). The key tasks were the fundamental steps in completing the operation, and these procedure specific skills technical skills were scored incrementally by 2 points in the likert scale depending how well the surgeon performs these skills. This was done after expert discussion, as the procedure specific technical skills were deemed more important to the generic technical skills, as they can affect patient outcome and quality of the operation.

Table 5.3 – Creation of Pneumoperitoneum

3	Create CO2 pneumoperitoneum	Do subtasks 3.1,3.2,3.3 in consecutive order	3.1 3.2 3.3	Open technique Insert 10mm umbilical port without trocar Insufflate abdomen with CO2	Pneumoperitoneum not created repeat 3.1 to 3.3 in consecutive order
---	-----------------------------	--	-------------------	--	---

Table 5.4 – Retraction & Exposure of Calot's Triangle

5	Laparoscopy, retraction, dissect & expose Calot's triangle (cystic artery, cystic duct)	Do subtasks 5.1, 5.2, 5.3 in consecutive order	5.1 5.2 5.3 5.4 5.5 5.6	Laparoscopy of abdomen Retract gall bladder Dissect adhesions to gallbladder Dissect & mobilise Hartmann's pouch Dissect & isolate cystic duct Dissect & mobilise cystic artery	Abnormal intra-abdominal pathology, decide to continue or not. Bleeding from port site – rectify before proceeding Graspers detach from gallbladder, reapply graspers Any significant bleeding during subtasks 5.3 to 5.6 will need correcting
---	---	--	--	--	--

Table 5.5 – Cystic Artery & Duct

7	Secure cystic artery & cystic duct	Do subtasks 7.1, 7.2, 7.3 in any order	7.1 7.2 7.3	Place 2 clips on proximal end of cystic artery Place clip on distal end of cystic artery Place clip at gallbladder end of cystic duct	Clips are not placed correctly or fall into abdomen. Retrieve & remove incorrect clips. Reapply clips 7.1 to 7.3 as necessary
---	------------------------------------	--	-------------------	---	---

Table 5.6 – Dissection of Gallbladder

12	Dissect gallbladder from liver bed	Do subtasks 12.1, 12.2, & 12.3 in any order then 12.4	12.1 12.2 12.3 12.4	Dissect left (medial) side of gallbladder up to fundus Dissect right (lateral) side of gallbladder up to fundus Separate under surface of gall bladder from liver Secure any bleeding from liver bed	Excessive bleeding, irrigate area and diathermy where necessary. Placement of Surgicel™ or similar material
----	------------------------------------	---	------------------------------	---	---

Table 5.7 – Removal of Gallbladder

13	Removal of dissected gallbladder	Do subtasks 13.1, 13.2, 13.3 in consecutive order	13.1 13.2 13.3	Insert retrieval bag Place gallbladder inside bag Extract bag containing gallbladder	Bag bursts, remove bag & repeat 13.1 to 13.3 in consecutive order. 10mm port too small for gallbladder and bag. Enlarge port site or remove gallstones from gallbladder before doing 13.3
----	----------------------------------	---	----------------------	--	---

Table 5.8 – Specific Likert Technical Skills Assessment Scale – Lap Chole

Specific Technical Skills	2	4	6	8	10
Access & Port Insertion	Created clumsily & with difficulty		Created adequately		Created quickly & skilfully
Gall Bladder Retraction & Exposure of Calot's triangle	Poor retraction & exposure of Calot's triangle		Satisfactory retraction & exposure of Calot's triangle		Expert retraction & exposure of Calot's triangle
Cystic Duct Dissection	Inadequate identification of cystic duct		Identified		Clearly identified
Cystic Duct Clipping & Transection	Clips not placed accurately. Bleeding		Clips placed proximally and distally adequately. Minimal bleeding		Clips placed expertly, proximally and distally, with adequate with length. No bleeding
Cystic Artery Dissection & Transection	Inadequate identification of cystic artery with clips not placed accurately.		Identified, clips proximally and distally adequate.		Clearly identified, clips proximally and distally adequate with enough length.
Gall Bladder Fossa Dissection	Gall bladder perforated during procedure with bile and stone spillage. Excessive Liver bed bleeding		Gall bladder only minimally perforated during procedure with no bile and stone spillage. Liver bed dry with minimal bleeding.		Gall bladder entirely intact during whole procedure. Liver bed dry with no bleeding.
Extraction of Gallbladder	Clumsily done with perforation of gallbladder		Adequately done with little trauma		Smoothly done with minimal trauma

Checklist Scale - Assessing Minor & Major Errors in Surgery

For this checklist scoring system we used an amalgamation of previous checklist and error scoring systems (Eubanks et al 1999, Seymour ² et al 2004, Tang ¹ et al 2004) in laparoscopic surgery as a template (Tables 5.9, 5.10, 5.11) as well as the literature search completed in Chapter 2. The scoring system (Table 5.12, 5.13, 5.14) is divided into three sections, after the refinement and feasibility study was done (see below). Minor events (errors) are inconsequential errors, major events are potential consequential errors if not corrected at the time, and significant major events are consequential errors which can affect patient outcome. We then adapted this scoring system for the open operations e.g. inguinal hernia repair and sapheno-femoral junction ligation. For the ***minor errors*** 0 points were awarded if no error was done, 4 points if done incorrectly but then rectified and 6 points if incorrectly done and not rectified. For ***major errors*** 0 points were awarded

if no error was done, 8 points if incorrectly being done but rectified and 12 points if incorrectly being done and not rectified. For *significant major errors* 0 points were awarded if no error was done, 16 points if incorrectly being done but rectified and 24 points if incorrectly being done and not rectified. These scores reflected the severity of the specific technical skills error on the quality of the operation and possible patient outcome.

Table 5.9 – Eubanks Error Scoring – Laparoscopic Cholecystectomy

Error	Error points	Frequency	Sum
<u>Gallbladder</u>			
Gallbladder Injury, Mechanical or Cautery (no bile spilled)	1		
Unintentional Release of the Gallbladder with Grasper	1		
Gallbladder Injury (bile or stones spilled)	10		
<u>Liver</u>			
Liver Injury (Including cautery) without Bleeding	1		
Liver Injury with Bleeding	5		
Major Vascular Injury (other than cystic artery)	50		
CBD, Hepatic Duct Injury	100		
<u>Cystic Duct</u>			
Additional Attempt at Clip/Ligature Placement on Duct	1		
Additional Attempt at Ductotomy	1		
Additional Attempt at Cystic Duct Cannulation	1		
Misplaced Clip or Ligature on Cystic Duct	2		
Unintentional Removal of Cholangiogram Catheter	5		
Unintentional Cystic Duct Transection	10		
Failure to Cannulate Patent Cystic Duct (only mark once)	10		
<u>Cystic Artery</u>			
Additional Attempt at Clip Placement on Artery	1		
Additional Attempt at Cutting Cystic Artery	1		
Misplaced Clip on Cystic Artery (clip on clip, partial occlusion, and so forth)	2		
Mistaking Artery for Duct (or Duct for Artery)	5		
Cystic Artery Tear	15		
<u>Miscellaneous</u>			
Injury to Other Abdominal Viscus	25		
Prolonged Operative Time (> 90 minutes, excluding cholangiogram)	10 pts/15 min		
			Total:

Table 5.10 – Seymour’s Error Scoring – Laparoscopic Cholecystectomy

Onset/offset Markers for Procedure Assessment:
1. Timing of length of procedure:
START - first contact of cautery with tissue
END – last attachment is divided
2. Time of general haemostatic inspection or waiting for an instrument may be subtracted.
Error Event Definitions
1. Lack of progress: No progress made in excising the gallbladder for an entire minute of the dissection. Dealing with the consequences of a predefined error represents lack of progress if no progress is made in excising the gallbladder during this period.
2. Gallbladder injury: There is gallbladder wall perforation with or without leakage of bile. Injury may be incurred with either hand.
3. Liver injury: There is liver capsule and parenchyma penetration, or capsule stripping with or without associated bleeding.
4. Incorrect plane of dissection: The dissection is conducted outside the recognized plane between the gallbladder and the liver (i.e., in the submucosal plane on the gallbladder, or subcapsular plane on the liver).
5. Burn nontarget tissue: Any application of electrocautery to nontarget tissue with the exception of the final part of the fundic dissection where some current transmission may occur.
6. Tearing tissue: Uncontrolled tearing of tissue with the dissecting or retracting instrument.
7. Instrument out of view: The dissecting instrument is placed outside the field of view of the telescope such that its tip is unviewable and can potentially be in contact with tissue. No error will be attributed to an incident of a dissecting instrument out of view as the result of a sudden telescope movement.
8. Attending takeover: The supervising attending surgeon takes the dissecting instrument (right hand), or retracting instrument (left hand) from the resident and performs a component of the procedure.

Table 5.11 – Tang's Error Scoring – Laparoscopic Cholecystectomy

Detail of Errors	Task			All Tasks
	1	2	3	
Consequential Errors				
Perforation				
Of the gallbladder	1	36	65	102
Of the small bowel	0	1	0	1
Bleeding				
From a liver injury	5	12	39	56
From a cystic artery	0	7	0	7
From a small vessels	11	60	20	91
From an omental injury	5	8	1	14
Injury to the cystic duct	0	11	0	11
Omitting to coagulate bleeding	1	4	0	5
Diathermy burn				
To the liver	9	134	243	386
To the diaphragm	0	0	5	5
To the duodenum	0	2	0	2
To the abdominal wall	0	0	2	2
Broken instrument	1	1	0	2
Total No. of Errors in Each Task	33	276	375	684
Inconsequential Errors				
Overshooting of instrument movement (use of excessive force with instrument traveling a longer distance than optimum)	19	158	154	331
Nonvisualization of instrument tip during dissection	27	209	95	331
Instrument movement out of endoscopic view	12	69	12	73
Applying current without visualizing instrument	1	16	7	24
Avulsion of the tissue rather than dissection	9	18	13	40
Cutting without lifting tissues from underlying structures	3	36	0	39
Inappropriate cutting with instruments	3	13	127	143
Insertion of instruments in wrong tissue planes	0	49	17	56
Inappropriate grasping of tissues	49	306	136	491
Total No. of Errors in Each Task	123	874	561	1558

*Task 1 indicates the division of adhesions involving the gallbladder and adjacent organs/omentum in the right upper quadrant; task 2, dissection of the cystic pedicle and division of the cystic artery and duct; and task 3, separation of the gallbladder from the liver bed followed by extraction.

Table 5.12 - Specific Technical Skills – Checklist Scoring - Minor Events

	Yes (U) 6	Yes (C) 4	OK 0		Yes (U) 6	Yes (C) 4	OK 0
Injury to gallbladder with bile spilled				Clip incompletely on cystic duct			
Liver injury, by diathermy				Misplaced clip fallen into abdomen			
Clip incompletely on cystic artery				Cystic artery or branches not identified initially			

Table 5.13 - Specific Technical Skills - Checklist Scoring - Major Events

	Yes (U) 12	Yes (C) 8	OK 0		Yes (U) 12	Yes (C) 8	OK 0
Gallbladder injury with stones spilled				Unintentional cystic duct division			
Liver injury with bleeding				Cystic artery injury			
Loss of pneumoperitoneum due to surgery							

Table 5.14 - Specific Technical Skills - Checklist Scoring – Significant Major Events

	Yes (U) 24	Yes (C) 16	OK 0		Yes (U) 24	Yes (C) 16	OK 0
Injury to other abdominal viscus				Duct Injury - CBD - Right Hepatic - Accessory			
Other major vascular injury							

(U) Yes Uncorrected

(C) Yes Corrected

OK = No error

5.13 Refinement & Feasibility of Scoring Systems

In this phase the two observers (experienced surgeons) sat together and watched 2 of each operation together (laparoscopic cholecystectomy, inguinal hernia, saphen-femoral junction ligation) on a large video screen. The video recordings of the operations were filmed using a Sony DCR-PC8E digital camcorder. Full length versions of operations were recorded onto digital video tape. These were then transferred without editing onto DVD, using Sony Click to DVD Software (Tokyo, Japan). Details of patients and subjects were recorded on computerized study forms, completed pre and postoperatively. They made any adjustments to the wording of the generic and specific technical skills

scoring systems after discussion. They used the generic and specific technical skills scoring systems for likert and checklist scoring, to assess if using such systems was practical.

5.14 Reliability of Scoring Systems

In a further test to evaluate the reliability of the refined scoring systems, the two observers (experienced surgeons) watched 3 of each operation independently and blindly (laparoscopic cholecystectomy, inguinal hernia, saphen-femoral junction ligation). Reliability studies using intra-class correlation and kappa coefficients for likert and checklist scoring respectively were calculated. The results of these reliability studies and mean assessment scores are shown in Table 5.15, 5.16, 5.17. The two observers showed good inter-rater reliability between them for all of the operations.

Table 5.15 – Individual Mean Scores & Inter-Rater Reliabilities Between Two Experienced Surgeons Laparoscopic Cholecystectomy

Category	Generic Likert	Specific Likert	Generic Checklist	Specific Minor Checklist	Specific Major Checklist	Specific Significant Major Checklist
Reliability	0.89	0.90	0.85	0.91	1.00	1.00
Observer 1	27	60	13.0	3.3	3	0
Observer 2	25	56	11.3	4.0	3	0

Table 5.16 – Individual Mean Scores & Inter-Rater Reliabilities Between Two Experienced Surgeons Open Inguinal Hernia Repair

Category	Generic Likert	Specific Likert	Generic Checklist	Specific Minor Checklist	Specific Major Checklist	Specific Significant Major Checklist
Reliability	0.92	0.94	0.89	1.00	1.00	1.00
Observer 1	26.0	43.3	12.0	0	0	0
Observer 2	28.0	41.0	11.3	0	0	0

Table 5.17 – Individual Mean Scores & Inter-Rater Reliabilities Between Two Experienced Surgeons Sapheno-Femoral Junction Ligation

Category	Generic Likert	Specific Likert	Generic Checklist	Specific Minor Checklist	Specific Major Checklist	Specific Significant Major Checklist
Reliability	0.82	0.89	0.97	0.94	1.00	1.00
Observer 1	27.7	31.0	13.0	1.3	0	0
Observer 2	24.0	27.7	12.7	2.7	0	0

5.15 Conclusions

In this chapter we have outlined the criteria required for a possible reliable and valid assessment of generic and specific technical skills tool. Also by using the hierarchical task analyses (Chapter 4) and the literature search of assessment of technical skills in surgery (Chapter 2), we have developed generic and specific likert and checklist scoring systems for technical skills in surgery for three key operations in General Surgery. The next three study chapters we will use these tools to assess open and laparoscopic live and simulated operations performed by Trainee and Consultant Surgeons, and demonstrate if the assessment tools have reliability and validity. Finally the assessment tool will be used to determine if novices can learn as a result of training in assessing generic and specific technical skills, and also determine what type of surgical experience is required to assess the operations accurately.

Chapter 6

Design of Studies

6.10 Overview

This chapter will briefly outline how the studies were designed and conducted. It will outline the design of the studies, the surgeons assessed, observers used, statistics used and finally the template for each study chapter.

6.11 Design

The main four prospective assessment study chapters in the thesis were conducted over a two year period from 2003 to 2004. The clinical studies were unrandomised, uncontrolled and they were based in the Greater and Central London area and also the south east of England. The studies were questionnaire and observer style based studies. Ethical approval was sought prospectively from the Health Authority, Surgeons and Patients.

6.12 Profile of Observing Assessors

Two main observing assessing surgeons had >12 years postgraduate surgical experience (Consultant Surgeon and Specialist Registrar) assessed each operation blindly and independently. The Consultant Surgeon had been a Consultant for more than 5 years. The Specialist Registrar had completed three years of Higher Surgical Training in the UK. Both surgeons had been involved in more than 150 laparoscopic cholecystectomies, primary inguinal hernia repairs and primary sapheno-femoral junction ligations.

6.13 Profile of Surgeons Assessed

The surgeons who were assessed were divided into two groups, Consultants Surgeons and Registrar Trainees. The Consultant Surgeons were General Surgeons in the Imperial

College network of hospitals. The trainees were Specialist Registrars in General Surgery in the North Thames West and South Thames West regions.

6.14 Statistics Used & Definitions

Reliability is a generic term that covers all aspects of the dependability of a measurement test or tool. For likert scoring intra-class correlation coefficient is used and for and checklist scoring kappa coefficient is used (Chapters 7,8,9).

Validity is the fundamental aspect of any assessment tool and that it measures or assesses what it claims to measure. Construct validity is the ability of an assessment tool to differentiate between experts and novices performing a given task or operation. Construct, concurrent, discriminate, and predictive validities are used in Chapters 7,8,9. Face and content validities are demonstrated in Chapters 4 and 5.

For the last study chapter (assessors of technical skills) the percentage agreement of each group of observers to the experienced surgeons was used. Corelation studies could not be done as there were too many observers in each assessor group.

6.15 Outline of Study Chapter Templates

Each study chapter will be set out with the following subheadings:

Overview, Methods & Materials, Results, Conclusions

Assessment Study

Chapter 7

Assessment of Technical Skills in Open Live Inguinal Hernia Surgery

7.10 Overview

As previously outlined assessing surgical technical skills in a structured manner is a topical issue in light of re-validation, shorter higher surgical training periods and recent high profile medico-legal cases.

To date there has been no systematic attempt to comprehensively assess both generic and procedure specific technical skills in live open and laparoscopic surgery. In this present study we aim to develop and validate new assessment tools which can assesses both generic and procedure specific technical skills in open live inguinal hernia surgery.

7.11 Methods & Materials

Design of Study

This was a prospective unrandomised observational clinical study of operations, conducted between 2003 to 2004 in London, UK. Patients, Trainees and Consultants Surgeons were recruited at specific elective lists at the ACAD theatre at Central Middlesex Hospital and the main theatre at St Mary's Hospital.

Procedures of Study

Hierarchical Task Analysis of Primary Inguinal Hernia Repair

A description of a primary inguinal hernia repair is given in the hierarchical task analysis chapter in the PhD. Hierarchical task analysis was used to identify the key tasks and sub-tasks in primary inguinal hernia repair, and was the basis to construct the assessment tool in assessing specific technical skills. Hierarchical task analysis was constructed for primary Lichtenstein inguinal hernia repair after referencing textbooks, articles, papers, web pages and expert panel discussions. The standard Lichtenstein method was used to repair the inguinal hernia, as this is widely used in the UK and worldwide, and is recommended by the Royal College of Surgeons of England. This operation was chosen as it is performed by Trainees and Consultants in large numbers, one of the commonest general surgical operations and is a fundamental operation to be learnt and performed by both Trainees and Consultants.

Development of the Assessment Tool

Likert scales were then constructed individually for generic and specific technical skills for the operation (Table 7.1A & 7.1B). Generic technical skills were scored incrementally by 1 point and procedure specific skills incrementally by 2 points. This was done after expert discussion, as the procedure specific technical skills were deemed more important to the generic technical skills as they can affect patient outcome and quality of the operation.

Checklists (Table 7.2A & 7.2B) were also constructed for generic and specific technical skills. In the later case, minor, major and significant major error scores were constructed and weighted after expert discussion. Generic technical skills were scored 0 point if correctly done and 1 point if incorrectly done. For minor errors, 0 points if no error was done, 4 points if done incorrectly but then rectified and 6 points if incorrectly done and not rectified. For major errors, 0 points for no error being done, 8 points if incorrectly being done but rectified and 12 points if incorrectly being done and not rectified. For significant major errors 0 points for no error being done, 16 points if incorrectly being done but rectified and 24 points if incorrectly being done and not rectified. These scores reflected the severity of the specific technical skills error on the quality of the operation and possible patient outcome.

Data Collection

The video recordings of the operations were filmed using a Sony DCR-PC8E digital camcorder. Full length versions of operations were recorded onto digital video tape. These were then transferred without editing onto DVD, using Sony Click to DVD Software (Tokyo, Japan). Details of patients and subjects were recorded on computerized study forms, completed pre and postoperatively.

Assessing the Operations

Two surgeons with >12 years postgraduate surgical experience (Consultant Surgeon and Specialist Registrar) assessed each operation blindly and independently. They were watched on large screen televisions and assessed using the likert and checklist technical

skills assessment tools. All of the operations were seen, and only slow parts (non-essential) of the operations were fast forwarded.

Patients & Setting

All 48 patients were between ASA grades 1-3, BMI < 32, < 75 years old and were from either the ACAD centre, Park Royal, London W5 or St Mary's Hospital, London W2. They were all discharged on the same or next day. There were no major post-operative complications.

Subjects

The surgeons were divided into two groups, Consultants (performed > 150 inguinal hernia repairs) and Registrar Trainees (performed < 150 inguinal hernia repairs). They were all explained about the study and verbally consented to participate in the study, and had the option not to participate. The video camera operator did not give advice on technical aspects of the operation before, during or after the operation to the surgeons. All the Registrar Trainee operations were unsupervised.

Ethical Approval

Local ethics committee approval was achieved and each patient had the opportunity for written consent and the surgeons had verbal consent. Either the patient or the surgeon had the opportunity to decline to be included in the study.

Statistical Analysis

Data was collated in an Excel[®] database (Microsoft, Redmond, Washington, USA). Statistical analysis was carried out with SPSS[®] software statistical package (SPSS, Chicago, Illinois, USA). $P < 0.05$ was regarded as statistically significant. For reliability analysis between the two observers two statistical methods were used. For likert scoring Intra-class Correlation Coefficient (ICC), also known as Cronbach α , and for checklists Kappa coefficients (k) were used. ICC and $k > 0.61$ was deemed statistically significant, with $p < 0.05$ (Appendix 7.3). The data for validity analysis was parametric for likert scores so the Repeated Measures Analysis of Variance (ANOVA) was used. For Checklists it was non-parametric and Mann-Whitney test was used for comparison between the two groups of surgeons (Consultants & Trainees).

For the time taken to do the operation for the two groups of surgeons, the data was parametric. A Wilcoxon test was used to compare the two groups. A senior statistician from the Statistical Advice Department at Imperial College London analysed the data for reliability and validity with the primary researcher.

7.12 Results

Descriptive Statistics & Assessments

There were 52 inguinal hernia repair operations recorded. Four had to be discarded due to video recording technical problems. 48 were then converted to DVD and analysed. A total of 16 surgeons participated in the study (Consultants = 10, Registrar Trainees = 6. 15 of whom were right-handed and one was left-handed. 13 were male surgeons and

three were female). Eight Consultants were less than five years in their present grade and two Consultants were more than five years as a Consultant Surgeon. There were two registrar trainees in years 5-6 and four trainees between years 1-4. Meaningful subgroup analysis of the Consultants and Trainees could not be done as each group did not have large numbers per surgeon. There were 38 Consultant operations, mean per operation 3.6 (range 1-10) and 10 Registrar Trainee operations, mean per operation 2 (range 1-4). Mean scores and ranges of the scores using likert and checklists of the two different grades of surgeons are shown in Table 7.3A &B.

Table 7.3A – Means & Ranges of Likert Scores of Surgeons – Open Hernia

Grade of Surgeon	Mean Likert Generic		Mean Likert Specific	
Consultants	25	Range 17-30	40	Range 18-50
Trainees	19	Range 15-24	32	Range 26-38

Table 7.3B – Means & Ranges of Checklist Scores of Surgeons – Open Hernia

Grade of Surgeon	Mean Checklist Generic		Mean Checklist Specific Minor		Mean Checklist Specific Major & Significant Major	
Consultants	0.29	Range 0-3	0.66	Range 0-4	2.21	Range 0-36
Trainees	0.90	Range 0-4	4.60	Range 0-14	1.20	Range 0-12

There is an overlap between Trainees and Consultants in the mean likert generic and specific scores as well as the checklists for generic, minor and major technical skills. Comparing the means for the likert and checklists scores between the groups of surgeons, it seems there is a larger numerical difference between the Consultants and Trainees using the likert than checklist scoring.

There were 8 minor errors which were corrected (4 by Consultants, 4 by Trainees) and there were 5 minor errors which were not corrected all done by Trainees. There was 1 major error which was corrected by the Consultants, and there were 7 major errors which were not corrected (6 by Consultants, 1 by Trainees). There were no significant major errors done by either group of surgeons. Full construct validity analysis will be done below, deciphering if there is a statistical difference between the Trainees and Consultants, and therefore examining the validation of the assessment tools. Regarding the patients there were 36 male and 12 female patients. Mean age was 59 (range 24-74 years) and mean BMI was 27.

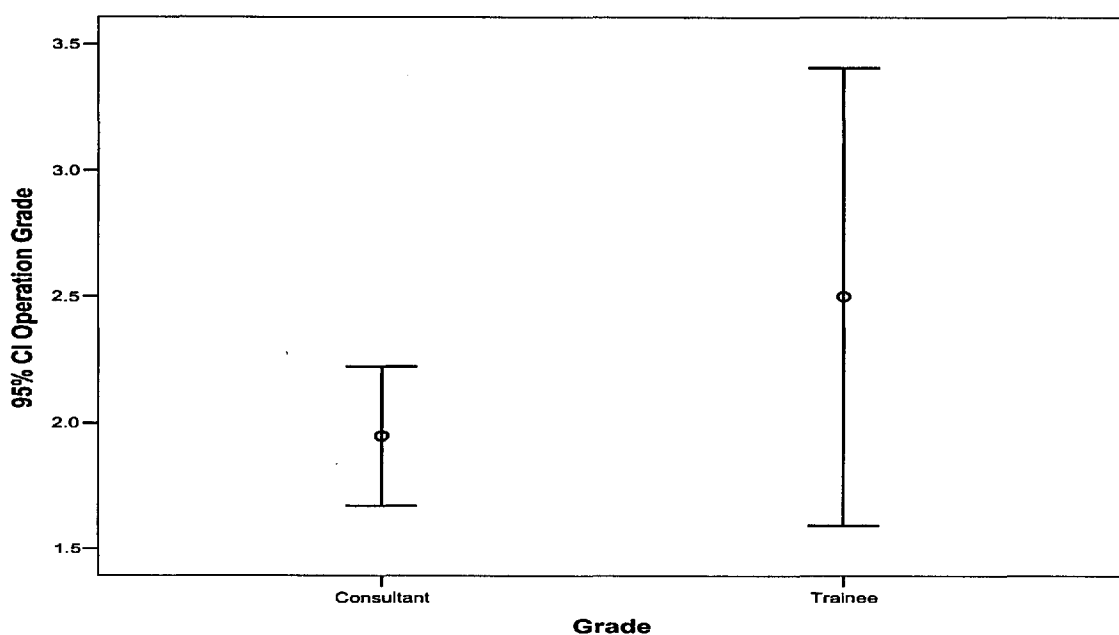
Grade of Operations

All of the operations were clinically graded into the following categories:

- 1/ Thin patient $BMI \leq 24$, direct hernia
- 2/ Thin patient $BMI \leq 24$, indirect hernia
- 3/ Normal/Moderate sized patient $BMI 25-29$, indirect or direct hernia
- 4/ Obese patient $BMI \geq 30$, direct or indirect hernia
- 5/ Inguinal-scrotal hernia

These were devised by the two assessors as these clinical aspects of the patient can affect the technical aspects of the operation. Figure 7.1 depicts mean operation grade and surgeon's level. Using non-parametric analysis, Chi-Square, there was no significant difference between the grade and level of surgeon, $p > 0.05$.

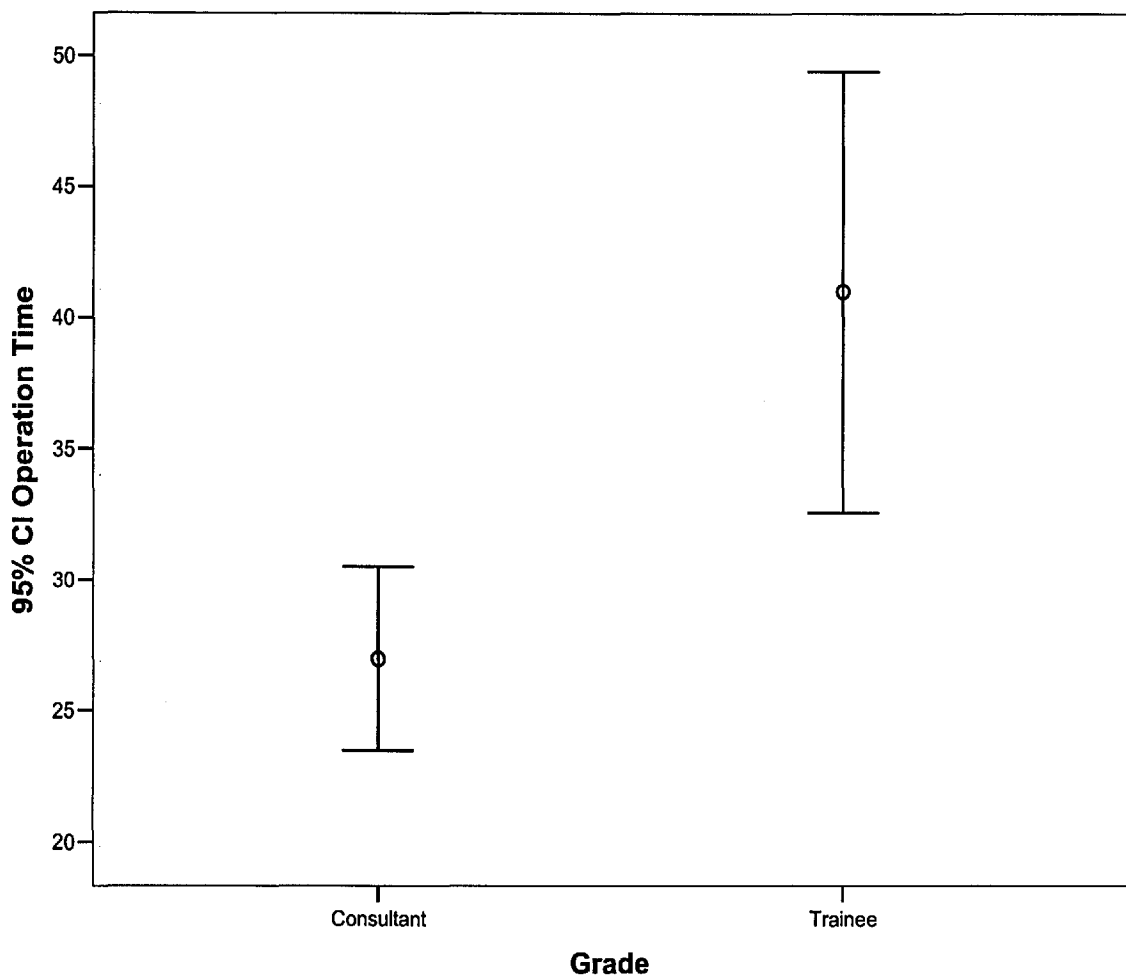
Figure 7.1 – Operation Grade & Level of Surgeon



Time Taken to Perform Operations

Mean time for the Consultant group to complete the operation was 27 minutes (range 15-75 minutes) and for the Registrar Trainee group was 41 minutes (range 20-55 minutes). This is displayed graphically in Figure 7.2. Using parametric analysis, Wilcoxon test, the difference was significant, $p < 0.05$.

Figure 7.2 – Operative Time & Level of Surgeon



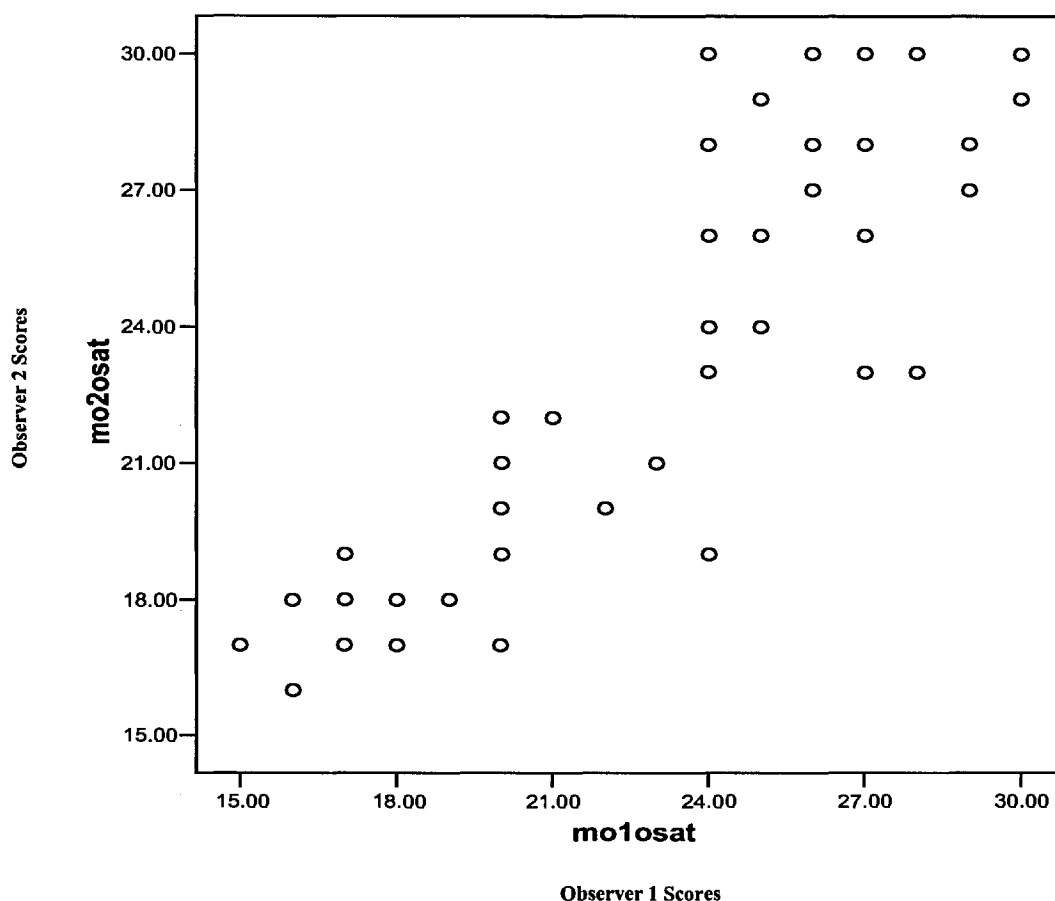
Likert Scoring Reliability – Between Observers

Generic Technical Skills - Likert Scoring

To assess reliability between the two observers for the likert technical assessment we use the intra-class correlation coefficient (ICC), also known as Cronbach α . Figure 7.3 depicts graphically the generic likert scores of observer 1 and 2. The Average Measure Intra-class Correlation ICC = 0.933, $p = 0.0000$. This means the reliability between the two observers is highly significant for generic likert scores.

Appendix 7.1 - for description

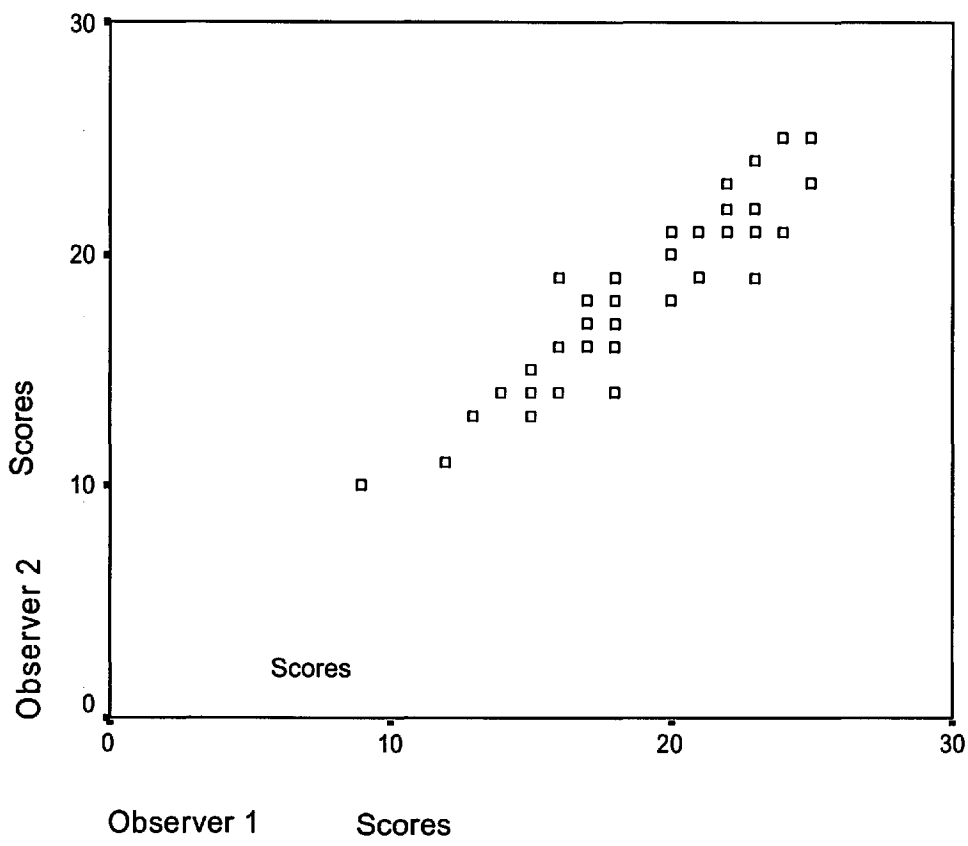
Figure 7.3 - Generic Likert Technical Skills Reliability Between Observers



Specific Technical Skills - Likert Scoring

The Average Measure Intra-class Correlation $ICC = 0.939$, $p = 0.0000$ for the two observers for the assessment of specific technical skills using likert scoring. Figure 7.4 depicts graphically the specific likert scores of observer 1 and 2. The statistics mean that the reliability between the two observers is highly significant.

Figure 7.4 - Specific Likert Technical Skills Reliability Between Observers



Checklist Scoring Reliability – Between Observers

To assess reliability between the two observers for the checklist technical assessment we use the weighted kappa coefficient (k). The reason why we use this method is explained in the appendices. Table 7.4 summaries the kappa coefficients for all the checklists used. It can be seen that all the checklists have a high reliability between the two observers.

Appendix 7.2 – for description

Appendix 7.3 – for description

Table 7.4 – Reliability of Technical Skills Checklists (Between Observers)

Technical Skills Being Assessed	Weighted Kappa Coefficients (k)	P Value	Conclusions
Generic Technical Skills	0.826	0.0000	Highly reliable between two observers
Specific Technical Skills Minor Events	0.871	0.0000	Highly reliable between two observers
Specific Skills Technical Major & Significant Major Events	0.993	0.0000	Highly reliable between two observers

Likert Technical Skills Validity – Between Consultants & Trainees

To assess validity for Trainees versus Consultants for the likert technical skills assessment we use Repeated Measures Analysis of Variance (ANOVA) using grade as a factor to test for differences between Trainees and Consultants.

Generic & Specific Technical Skills – Likert Rating

Figure 7.5 and 7.6 depicts graphically the statistical analysis. Using Repeated Measures Analysis of Variance (ANOVA) for construct validity shows that there is a significant difference between Trainees and Consultants in both generic (F-value = 30.159) and specific (F-value = 15.124) technical skills, $p < 0.05$. This means that on average Consultants perform generic and specific technical skills better than Trainees.

Figure 7.5 - Generic Technical Skills - Likert Scoring

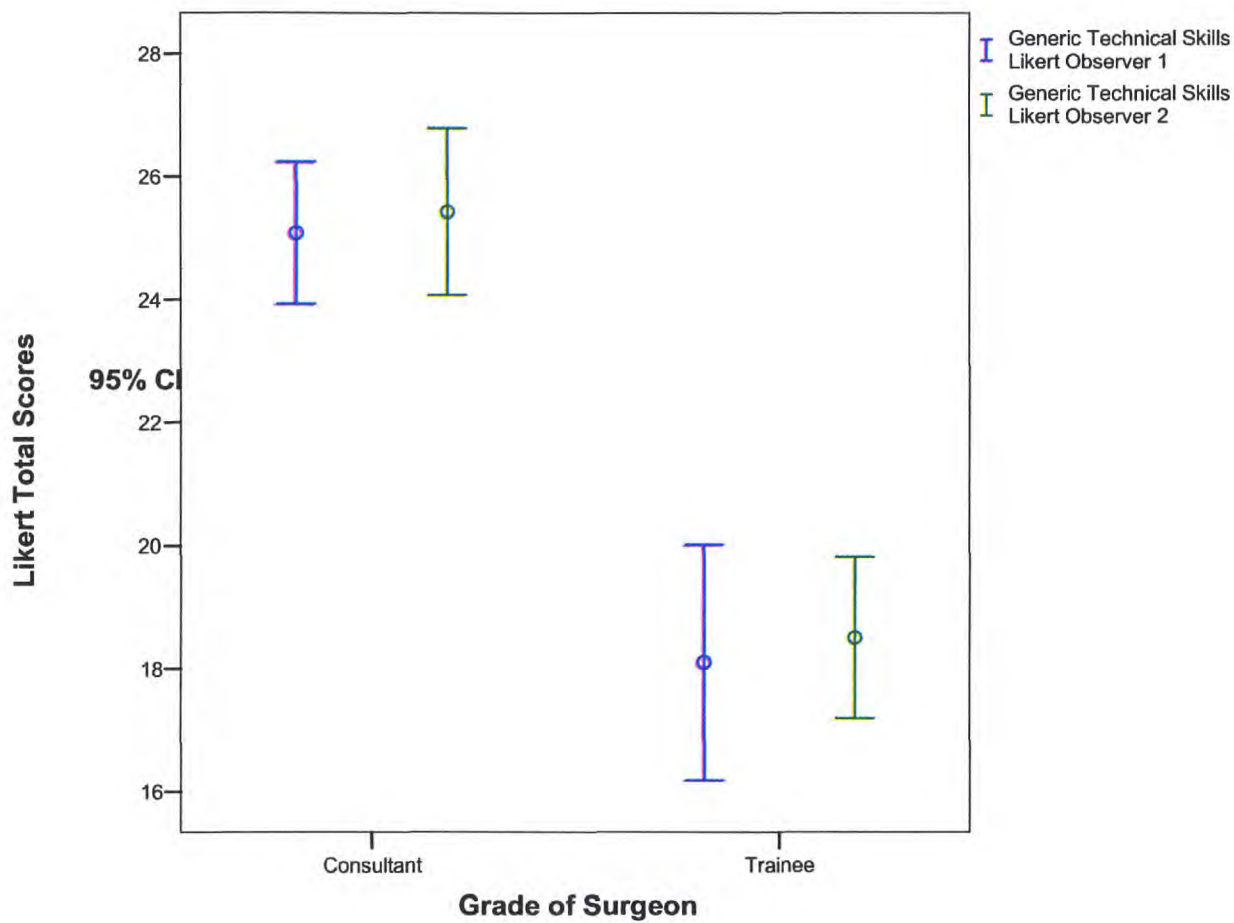
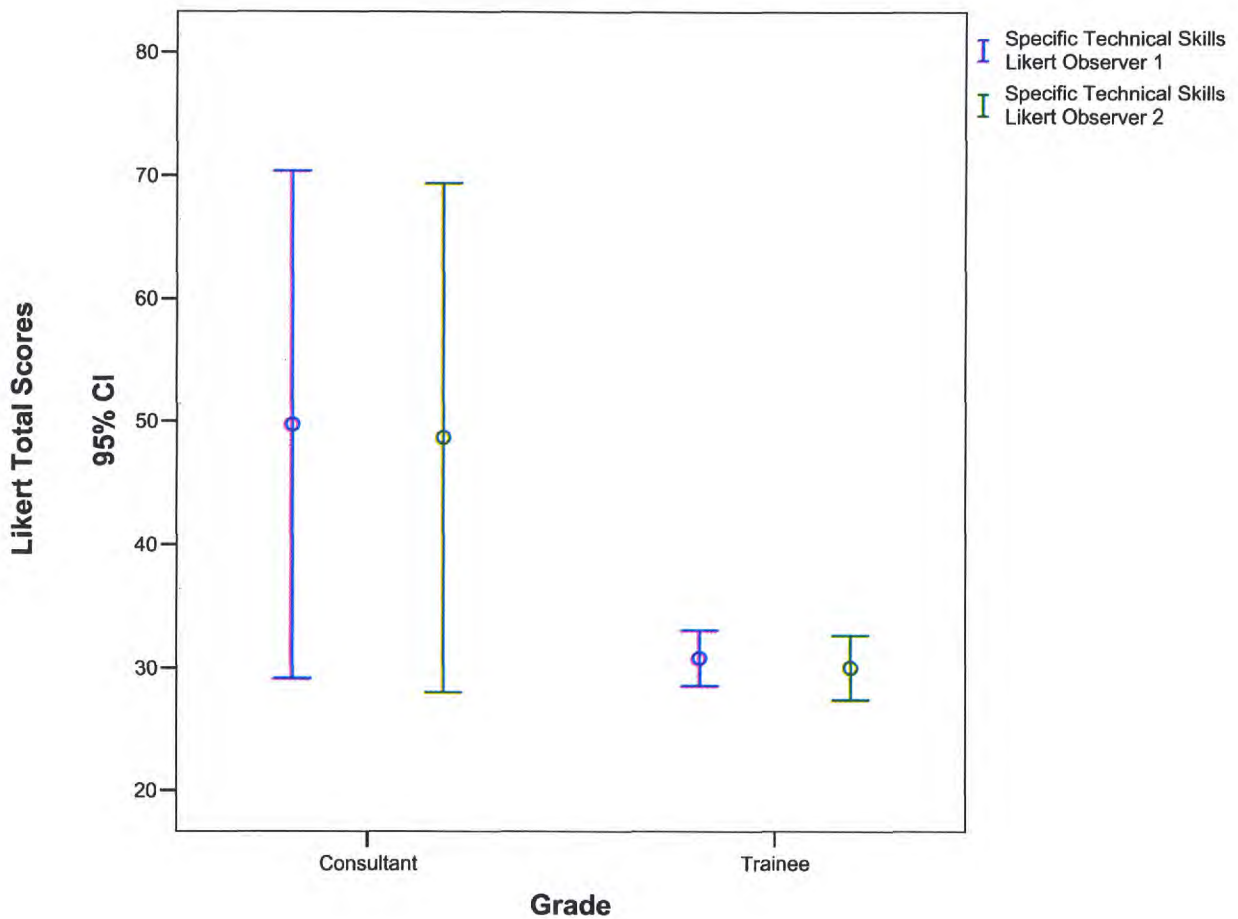


Figure 7.6 - Specific Technical Skills – Likert Scoring



Checklist Technical Skills Validity – Between Consultants & Trainees

We are going to apply different techniques from the ones used in the likert scoring assessment for the reason mentioned in the reliability analysis, because the variables can not be considered continuous. We can test for differences between Consultants and Trainees through a non-parametric test for two independent samples i.e. Mann-Whitney using the average of the scores. It is useful to display a graph, such as error bar charts, with the central number being the mean checklist score with a 95% confidence interval (Figure 7.7, 7.8, 7.9).

Lower scores relate to fewer errors in the generic technical skills checklist scale. In the minor, major and significant major events checklist (procedure specific technical skills), the **higher scores relate to more errors** being executed. Table 7.5 summaries the Mann-Whitney test results for the construct validity of each. There were no significant major errors done by both Trainees and Consultants, so the scores for major and significant major error scores were combined.

Table 7.5 – Validity of Technical Skills Checklists Between Consultants & Trainees

Technical Skill being Assessed	P Value	Conclusions
Generic Technical Skills	0.0072	Significant difference between Trainees & Consultants
Specific Technical Skills Minor Events	0.0002	Significant difference between Trainees & Consultants
Specific Skills Technical Major & Significant Major Events	0.3314	No difference between Trainees & Consultants

Figure 7.7 – Generic Technical Checklist Scores (Consultants & Trainees)

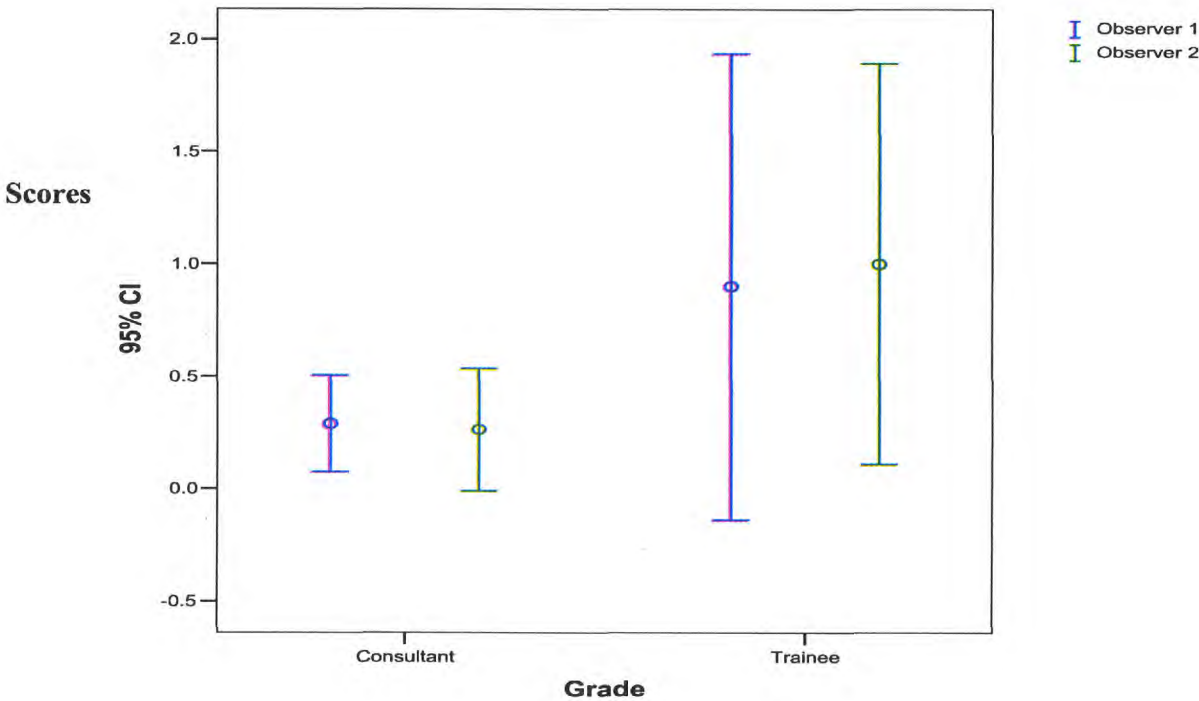


Figure 7.8 - Specific Minor Technical Checklist Scores (Consultants & Trainees)

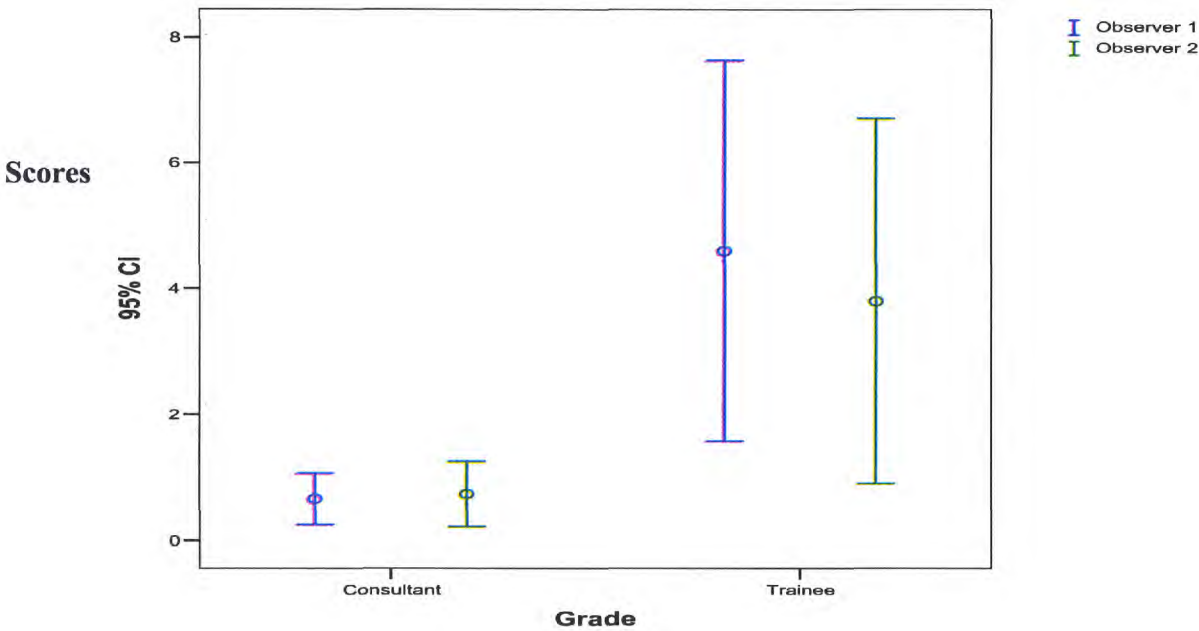
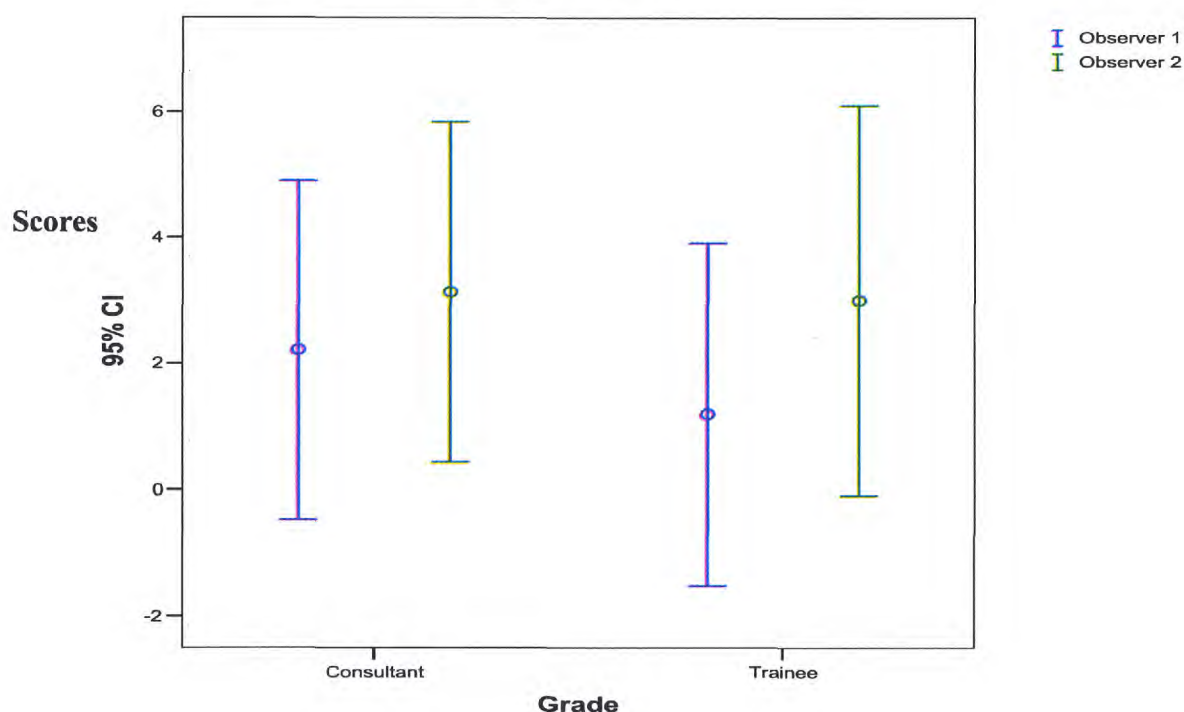


Figure 7.9 - Specific Technical Skills Checklist Major & Significant Major Scores (Consultants & Trainees)



7.12 Conclusions

Reliability for the two observers was highly significant for the both likert and checklists scoring for generic and specific technical skills. Construct validity analysis for the likert scoring generic and specific assessment of technical skills were both significant. As a group Consultants performed the operations better compared to Trainees regarding generic and specific technical skills using likert scoring.

Construct validity analysis for the checklist generic and specific minor technical skills were significant respectively. Consultants as a group performed the operations better compared to Trainees regarding generic and specific technical skills (minor) using the

checklist scoring. However for the procedure specific technical skills (major and significant major) checklist construct validity was not significant between the two groups.

The comparison of the operating times between the Consultants Surgeons and the Trainees was significant. This is due to that fact that as a surgeon performs more operations his or her generic technical skills become more proficient and hence faster. As a result the process of executing each task and sub-task becomes significantly faster. The Consultants had performed more than 150 primary inguinal hernia repairs; the Trainees had performed less than 150 operations. This assessment tool of technical skill in open live surgery seems to have face, content, concurrent, predictive and construct validities in the likert scales, and for the checklists for generic and specific technical skills (minor events). It also demonstrates good reliability in the assessment of likert and checklists between the two observers.

However comparing likert and checklist scoring it seems there is a larger numerical difference between Trainees and Consultants using the likert scoring than the checklist scoring. This may reflect that the likert scoring system is more sensitive, as checklist scoring is finite and is unable to accommodate the subtle range of technical skills in various grades of surgeons. All of the surgeons in the study were deemed technically competent, hence they were not supervised by a senior surgeon, and therefore no group difference appeared in the specific technical skills checklist (major and significant major) categories. These categories would have an affect on patient outcome and quality of the final operation.

Table 7.1A – Generic Technical Skills Likert Scale for Inguinal Hernia Repair

Category	1	2	3	4	5
Handling & Dissection of Tissues	Unnecessary force on tissues, causing frequent damage to tissues		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time & Motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument & Material Handling	Repeatedly makes tentative or awkward moves with instruments or material		Competent use of instruments/material although occasionally awkward		Fluid moves with instruments/material and no awkwardness
Use of Instruments & Materials	Frequently used inappropriate instruments & materials		Used appropriate instrument & materials for the task or subtask		Expertly used appropriate instruments & materials
Exposure of Technical Field	Consistently poor exposure		Good exposure most of the time		Expert exposure
Completion & Flow of Tasks & Subtasks	Frequently stopped operating or hesitant in next task or subtask		Forward planning with steady progression of tasks & subtasks		Effortless flow from one task to subtask

Table 7.1B – Specific Technical Skills Likert Scale for Inguinal Hernia Repair

Specific Skill	2	4	6	8	10
Exposure of external oblique aponeurosis & external ring	Not clearly identified & exposed		Competently identified & exposed		Expertly identified & exposed
Mobilisation of spermatic cord	Not clearly identified & injury to spermatic cord		Competently identified & no injury to spermatic cord		Expertly identified & preserved
Exposure of hernia sac	Not clearly identified & excessive bleeding		Competently identified & minimal bleeding		Expertly identified & no bleeding
Management of direct/indirect hernial sac	Direct hernia – poorly plicated posterior wall Indirect hernia – poorly opened, reduced & suture ligated hernial sac. Viscera injured		Direct hernia – competently plicated posterior wall Indirect hernia – competently opened, reduced & suture ligated hernial sac. No visceral damage		Direct hernia – expertly plicated posterior wall Indirect hernia – expertly opened, reduced & suture ligated hernial sac. No visceral damage.
Insertion of mesh	Poorly cut & secured mesh and not medial to pubic tubercle, along inguinal ligament and around internal ring		Competently cut & secured mesh medial to pubic tubercle, along inguinal ligament and around internal ring		Expertly cut & secured mesh medial to pubic tubercle, along inguinal ligament and around internal ring

Table 7.2A – Generic Technical Skills Checklists Scale for Inguinal Hernia Repair

Category	0	1
	Yes	No
Appropriate access to technical field		
Correct understanding of tissue planes		
Demonstrates sound knowledge of anatomy		
Confident & efficient progression of tasks & subtasks		
Uses instruments & materials efficiently & correctly		
Appropriate exposure/retraction of technical field		
Possesses good hand-eye co-ordination		
Handles tissues gently with minimal damage		
Correct suturing and knotting techniques		
Uses diathermy appropriately and safely		
Able to control bleeding efficiently & competently		
Appropriate wound management - closure & dressings		
Tasks or subtasks not rushed or too slow		

Table 7.2B – Specific Technical Skills Checklists Scale for Inguinal Hernia Repair

(a) Specific Technical Skills – Checklist Scoring - Minor Events

	Yes (U) 6	Yes (C) 4	OK 0		Yes (U) 6	Yes (C) 4	OK 0
Inappropriate site of incision				Sac inadvertently breached during dissection			
Ilio-inguinal nerve cut unintentionally				Incomplete dissection of internal ring			

(b) Specific Technical Skills - Checklist Scoring - Major Events

	Yes (U) 12	Yes (C) 8	OK 0		Yes (U) 12	Yes (C) 8	OK 0
Poorly cut mesh				New internal ring not narrowed			
Mesh not secured medial to pubic tubercle & along inguinal ligament				External oblique aponeurosis not adequately closed			

(c) Specific Technical Skills - Checklist Scoring – Significant Major Events

	Yes (U) 24	Yes (C) 16	OK 0		Yes (U) 24	Yes (C) 16	OK 0
Damage to spermatic cord				Damage to contents of sac/viscera			

(U) Yes Uncorrected

(C) Yes Corrected

OK = No error

Assessment Study

Chapter 8

Assessment of Technical Skills in Open Live & Simulated Sapheno-Femoral Surgery

8.10 Overview

As previously outlined assessing surgical technical skills in a structured manner is a topical issue. With new synthetic innovations, it is possible to replicate some procedures in bench models. To validate these models, they must demonstrate face and content validities as well as construct validity (able to discriminate experts from novices).

To date there has been no tool to assess both generic and procedure specific technical skills in live and simulated open surgery. In this present study we aim to develop and validate new assessment tools which can assess both generic and procedure specific technical skills in open live and simulated sapheno-femoral (SF) surgery.

8.11 Methods & Materials

Design of Study – Live Open SF Surgery

This was a prospective unrandomised observational clinical study of operations, conducted between 2003 to 2004 in London, UK. Patients, Trainees and Consultants Surgeons were recruited at specific elective lists at the ACAD theatre at Central Middlesex Hospital.

Design of Study – Simulated Open SF Surgery

This was a prospective unrandomised non-clinical study, conducted on a specified day in July 2003 at Frimley Park Hospital, Surrey, UK. Registrar Trainees (Years 1-3) were recruited from the South West Thames region.

Procedures of Study

Hierarchical Task Analysis of Sapheno-Femoral Junction Ligation

A description of a primary sapheno-femoral junction ligation is given in the previous hierarchical task analysis chapter in the PhD. Hierarchical task analysis was used to identify the key tasks and sub-tasks in sapheno-femoral junction ligation, and was the basis to construct the assessment tool in assessing specific technical skills. Hierarchical task analysis was constructed for sapheno-femoral junction ligation after referencing textbooks, articles, papers, web pages and expert panel discussions. This operation was chosen as it is performed by Trainees and Consultants in large numbers, one of the commonest general vascular surgical operations and is a fundamental operation to be learnt and performed by both Trainees and Consultants in General Surgery.

Development of the Assessment Tool

Likert scales were then constructed individually for generic and specific technical skills for the operation (Table 8.1A & 8.1B). Generic technical skills were scored incrementally by 1 point and procedure specific skills incrementally by 2 points. This was done after expert discussion, as the procedure specific technical skills were deemed more important to the generic technical skills as they can affect patient outcome and quality of the

operation. Checklists (Table 8.2A & 8.2B) were also constructed for generic technical skills and specific technical skills. In the later case, minor, major and significant major error scores were constructed and weighted after expert discussion. If no generic technical skills error was done this was scored 0 point, if there was an error this was scored 1 point. For minor errors, 0 points if no error was done, 4 points if done incorrectly but then rectified and 6 points if incorrectly done and not rectified. For major errors, 0 points if no error was done, 8 points if incorrectly being done but rectified and 12 points if incorrectly being done and not rectified. For significant major errors, 0 points if no error was done, 16 points if incorrectly being done but rectified and 24 points if incorrectly being done and not rectified. These scores reflected the severity of the specific technical skills error on the quality of the operation and possible patient outcome.

Data Collection

The video recordings of the operations were filmed using a Sony DCR-PC8E digital camcorder. Full length versions of operations were recorded onto digital video tape. These were then transferred without editing onto DVD, using Sony Click to DVD Software (Tokyo, Japan). Details of patients and subjects were recorded on computerized study forms, completed pre and postoperatively.

Assessing the Operations

Two surgeons with >12 years postgraduate surgical experience (Consultant Surgeon and Specialist Registrar) assessed each live operation blindly and independently. For the bench models the Specialist Registrar solely assessed the simulated operations.

They were watched on large screen televisions and assessed using the likert and checklist technical skills assessment tool. All of the operations were seen, and only slow parts (non-essential) of the operations were fast forwarded.

Patients & Settings

All 27 patients were between ASA grades 1-3, BMI < 32, < 75 years old and were from the ACAD centre, Park Royal, London W5. They were all discharged on the same or next day. There were no major post-operative complications.

Bench Models & Settings

20 Sapheno-Femoral Junction Ligation bench models developed by Limbs & Things™ (Bristol, UK) were purchased. This model has previously been validated for face and content validities. The bench model operations were performed in the Clinical Skills Laboratory at Frimley Park Hospital.

Subjects

The surgeons were divided into two groups, Consultants (performed > 150 Sapheno-Femoral Junction Ligation) and Registrar Trainees (performed < 150 Sapheno-Femoral Junction Ligation). They were all explained about the study and verbally consented to participate in the study, and had the option not to participate. The video camera operator did not give advice on technical aspects of the operation before, during or after the operation to the surgeons. All the Registrar Trainee live and simulated operations were unsupervised.

Ethical Approval

Local ethics committee approval was achieved and each patient had the opportunity for written consent and the surgeons had verbal consent. Either the patient or the surgeon had the opportunity to decline to be included in the study.

Statistical Analysis

Data was collated in an Excel[®] database (Microsoft, Redmond, Washington, USA). Statistical analysis was carried out with SPSS[®] software statistical package (SPSS, Chicago, Illinois, USA). $P < 0.05$ was regarded as statistically significant.

For reliability analysis between the two observers two statistical methods were used. For likert scoring Intra-class Correlation Coefficient (ICC), also known as Cronbach α , and for checklists Kappa coefficients (k) were used. ICC and $k > 0.61$ was deemed statistically significant, with $p < 0.05$ (Appendix 7.3). The data for validity analysis was parametric for likert scores so the Repeated Measures Analysis of Variance (ANOVA) was used. For checklists it was non-parametric and Mann-Whitney test was used for comparison between the two groups of surgeons.

For the time taken to do the operation for the two groups of surgeons, the data was parametric. Wilcoxon test was used to compare the two groups. A senior statistician from the Statistical Advice Department at Imperial College London analysed the data for reliability and validity with the primary researcher. The live and simulated operations

were combined for statistical analysis as there were insufficient Trainees in the live operation group alone.

8.12 Results

Descriptive Statistics & Assessments – Live Open SF Surgery

There were 32 live open sapheno-femoral junction ligation operations recorded. Five had to be discarded due to video recording technical problems. 27 were then converted to DVD and analysed. A total of 7 surgeons participated in the study (Consultants = 4, Registrar Trainees = 3. Seven of whom were right-handed and none were left-handed. 7 were male surgeons and none were female). 2 Consultants were less than five years in their present grade and 2 Consultants were more than five years as a Consultant Surgeon. There were 3 Registrar Trainees in years 5-6 and no trainees between years 1-4. There were 22 Consultant operations, mean per operation 5.5 (range 1-12) and 5 Registrar Trainee operations, mean per operation 1.7 (range 1-2). Mean scores and ranges of the scores using likert and checklists of the two different grades of surgeons are shown in Table 8.3A & B.

Descriptive Statistics & Assessments – Simulated Open SF Surgery

There were 20 simulated open sapheno-femoral junction ligation operations recorded. They were then converted to DVD and analysed. A total of 20 Trainee Surgeons participated in the study. 18 of who were right-handed and 2 were left-handed. 17 were male surgeons and 3 were female). All Trainees were between years 1-3. Each Trainee

performed one operation each. Mean scores and ranges of the scores using likert and checklists of the two different grades of surgeons are shown in Table 8.4A &B.

Table 8.3A – Means & Ranges of Likert Scores of Surgeons – Live Open SF

Grade of Surgeon	Mean Likert Generic		Mean Likert Specific	
Consultants	27.1	Range 17-30	33	Range 18-50
Trainees	24.8	Range 15-24	34.4	Range 26-38

Table 8.3B – Means & Ranges of Checklist Scores of Surgeons– Live Open SF

Grade of Surgeon	Mean Checklist Generic		Mean Checklist Minor		Mean Checklist Major & Significant Major	
Consultants	0.09	Range 0-1	0.73	Range 0-4	7.63	Range 0-24
Trainees	0	Range 0	0.80	Range 0-4	0	Range 0

Regarding the patients there were 7 male and 20 female patients. Mean age was 54 (range 28-72 years) and mean BMI was 28.

There is an overlap between Trainees and Consultants in the mean likert generic and specific scores as well as the checklists for generic, minor and major technical skills in the live operations.

In the live operations there was only one generic technical skills error done by a Consultant. There were 5 minor errors which were corrected (4 by Consultants, 1 by Trainees) and there were no minor errors which were not corrected. There were no major errors done by either group of surgeons. There were 3 significant major errors which were corrected, and there were 5 significant major errors which were not corrected, both were all performed by Consultants.

Table 8.4A – Means & Ranges of Likert Scores of Surgeons – Simulated Open SF

Grade of Surgeon	Mean Likert Generic	Mean Likert Specific
Trainees	19.9 Range 13-25	22.9 Range 14-32

Table 8.4B – Means & Ranges of Checklist Scores of Surgeons – Simulated Open SF

Grade of Surgeon	Mean Checklist Generic	Mean Checklist Minor	Mean Checklist Major & Significant Major
Trainees	0.65 Range 0-6	0.60 Range 0-4	7.20 Range 0-48

In the simulated operations there were 13 generic technical skills error done by the Trainees. There were 3 minor errors which were corrected and there were no minor errors which were not corrected. There was 1 major error not corrected and 2 major errors not corrected. There was 1 significant major error which was corrected, and there were 4 significant major errors which were not corrected. Overall it seems that the Trainees in

the simulated open SF group perform more technical errors (all types) compared to the Consultants and Trainees in the live open SF group.

Full construct validity analysis will be done below, deciphering if there is a statistical difference between the Trainees and Consultants, and therefore examining the validation of the assessment tools.

Grade of Operations – Live Open SF

All of the operations were clinically graded into the following categories:

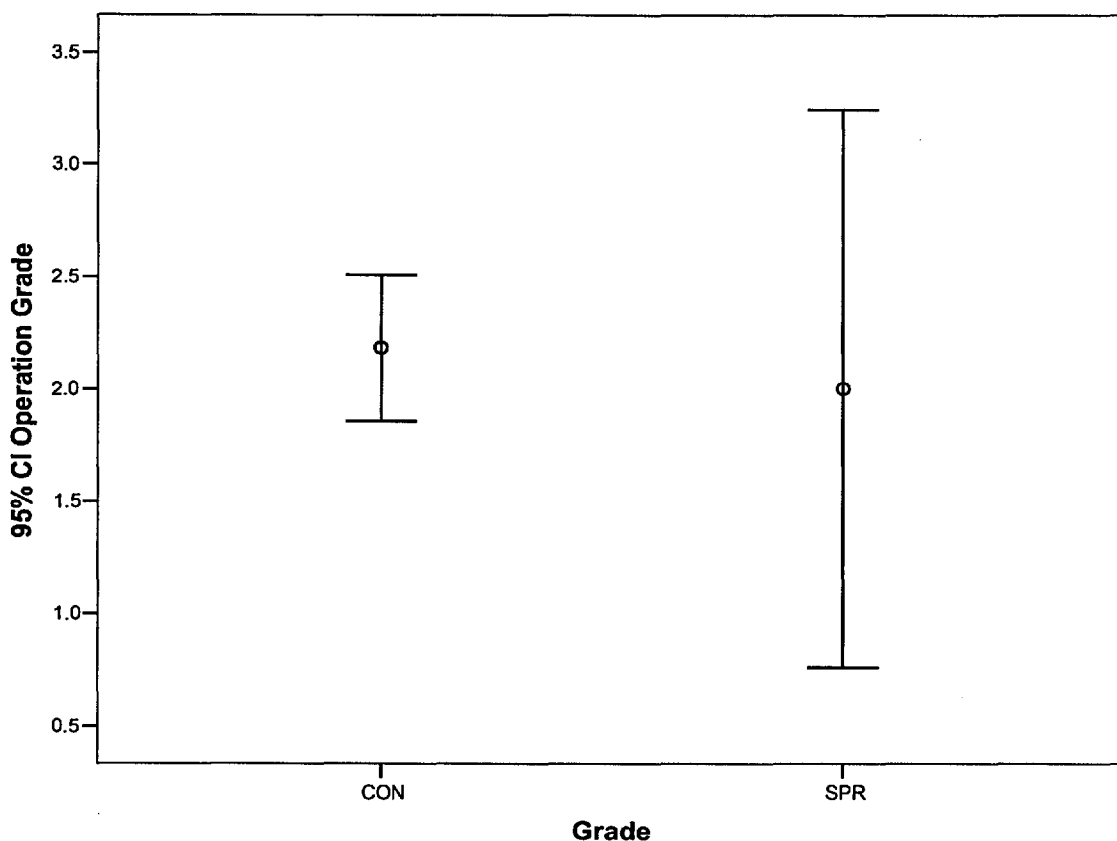
1/ Thin patient BMI ≤ 24

2/ Normal/Moderate sized patient BMI 25-29

3/ Obese patient BMI ≥ 30

These were devised by the two assessors as these clinical aspects of the patient can affect the technical aspects of the operation. Figure 8.1 depicts mean operation grade and surgeon's level. Chi-Square test did not demonstrate a difference between operation grade and surgeon's grade, $p > 0.05$.

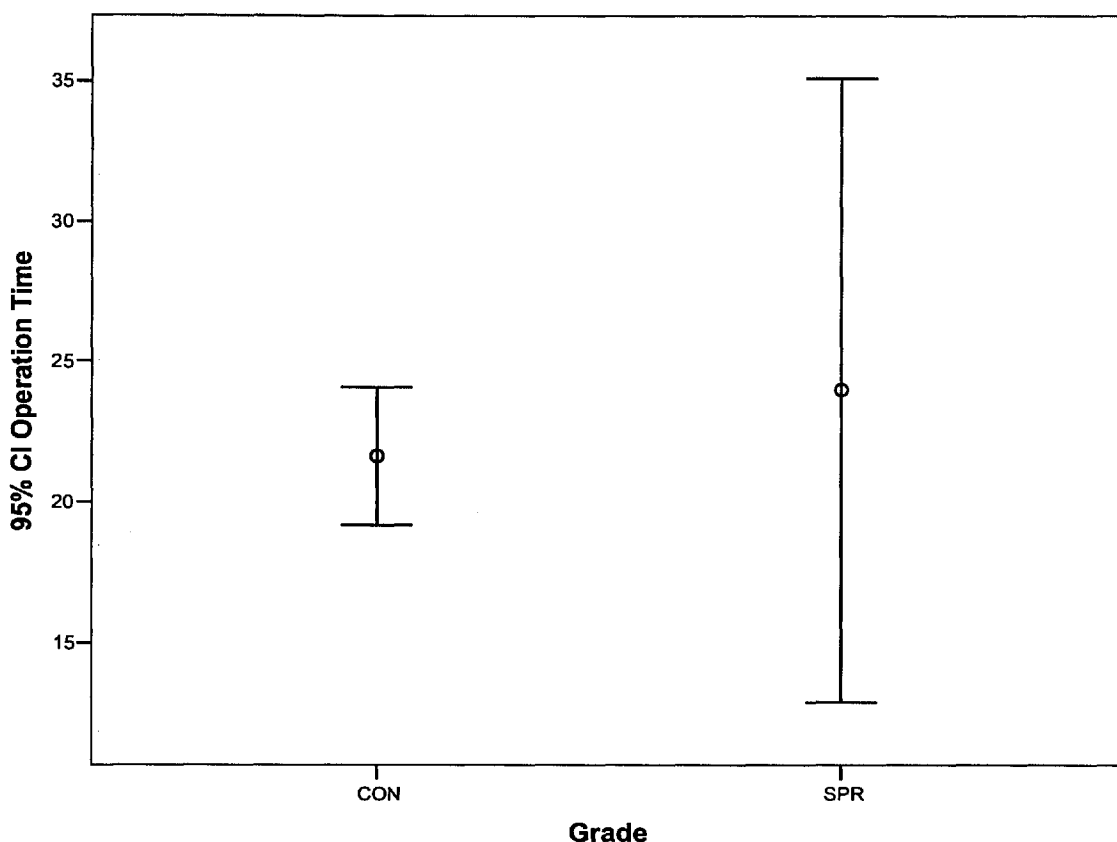
Figure 8.1 – Operation Grade & Level of Surgeon – Live Open SF



Time Taken to Perform Operations – Live Open SF

Mean time (Figure 8.2) for the Consultant group to complete the live operation was 22 minutes (range 19-24 minutes) and for the Trainee group was 24 minutes (range 13-35 minutes). Using parametric analysis, Wilcoxon test, the difference was not significant, $p > 0.05$.

Figure 8.2 – Operative Time & Level of Surgeon – Live Open SF



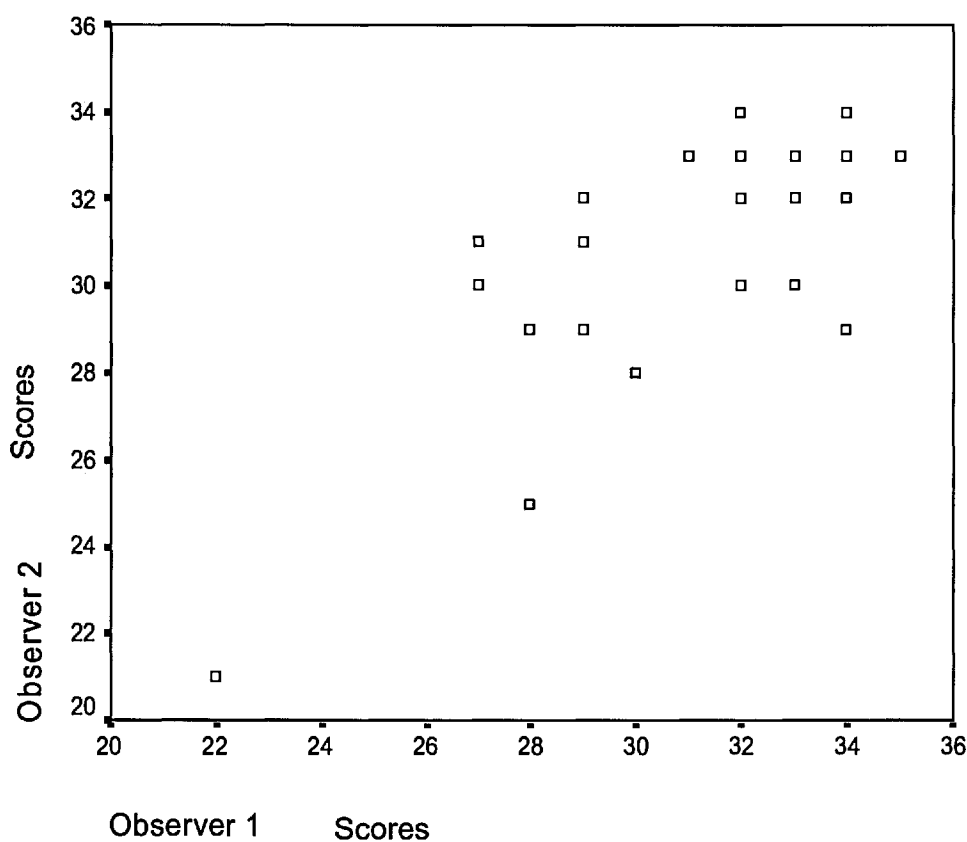
Likert Scoring Reliability – Between Observers for Live Open SF

Likert Technical Skills - Likert Scoring - Live Open SF

To assess reliability between the two observers for the likert technical assessment we use the intra-class correlation coefficient (ICC), also known as Cronbach α . Figure 8.3 depicts graphically the generic likert scores of observer 1 and 2. The Average Measure Intra-class Correlation ICC = 0.743, $p = 0.0004$. This means the reliability between the two observers is significant for generic likert scores.

Appendix 7.1 - for description

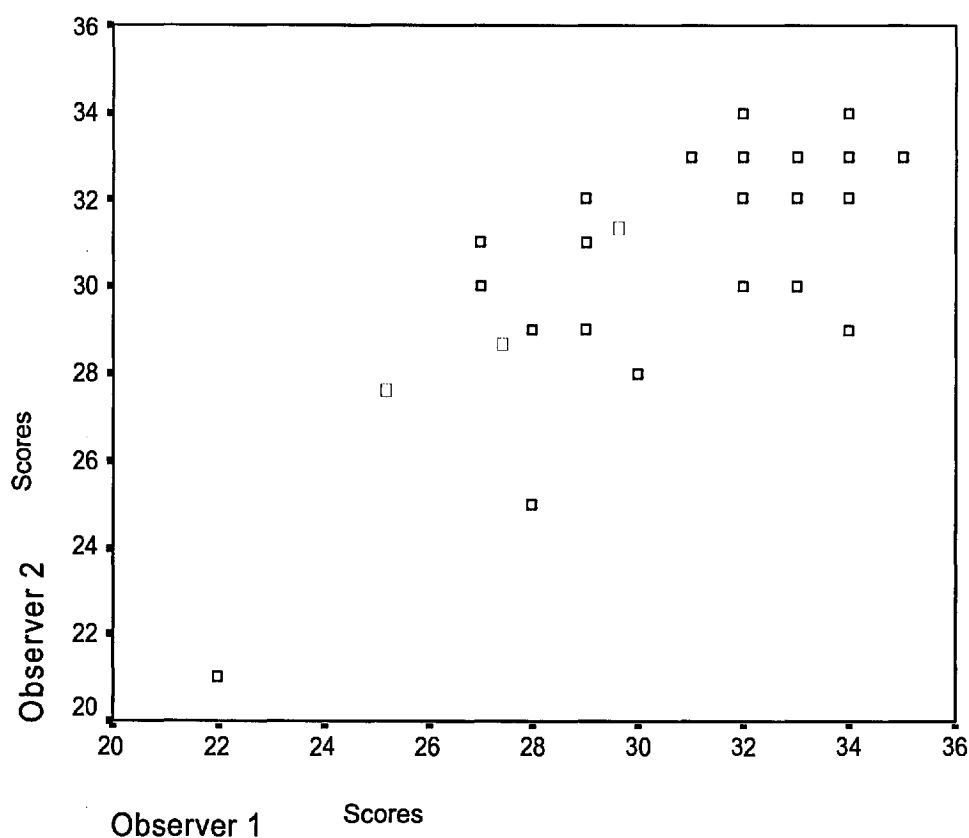
Figure 8.3 - Generic Likert Technical Skills Reliability – Live Open SF



Specific Technical Skills - Likert Scoring - Live Open SF

The Average Measure Intra-class Correlation $ICC = 0.712$, $p = 0.005$ for the two observers for the assessment of specific technical skills using likert scoring. Figure 8.4 depicts graphically the specific likert scores of observer 1 and 2. The statistics mean that the reliability between the two observers is significant.

Figure 8.4 - Specific Likert Technical Skills Reliability – Live Open SF



Checklist Scoring Reliability – Between Observers for Live Open SF

To assess reliability between the two observers for the checklist technical assessment we use the weighted kappa coefficient (k). The reason why we use this method is explained in the following appendices. Table 8.5 summaries the kappa coefficients for all the checklists used. It can be seen that all the checklists have a high reliability between the two observers.

Appendix 7.2 – for description

Appendix 7.3 – for description

Table 8.5 – Reliability of Technical Skills Checklists – Live Open SF

Technical Skill being Assessed	Weighted Kappa Coefficients (k)	P Value	Conclusions
Generic Technical Skills	1.00	0.0000	Highly reliable between two observers
Specific Technical Skills Minor Events	0.886	0.0000	Highly reliable between two observers
Specific Skills Technical Major & Significant Major Events	0.950	0.0000	Highly reliable between two observers

Likert Rating Validity – Between Consultants & Trainees – Live Open SF

Generic & Specific Technical Skills - Live Open SF

Using Repeated Measures Analysis of Variance (ANOVA) for construct validity (Figures 8.5 & 8.6) shows that there is no significant difference between Trainees and Consultants in generic ($F = 1.178$, $p = 0.288$) and specific technical skills ($F = 5.865$, $p = 0.231$). This means that on average Consultants perform generic and specific technical skills same as the Trainees.

Figure 8.5 – Generic Likert Technical Skills Scoring of Consultants & Trainees – Live Open SF

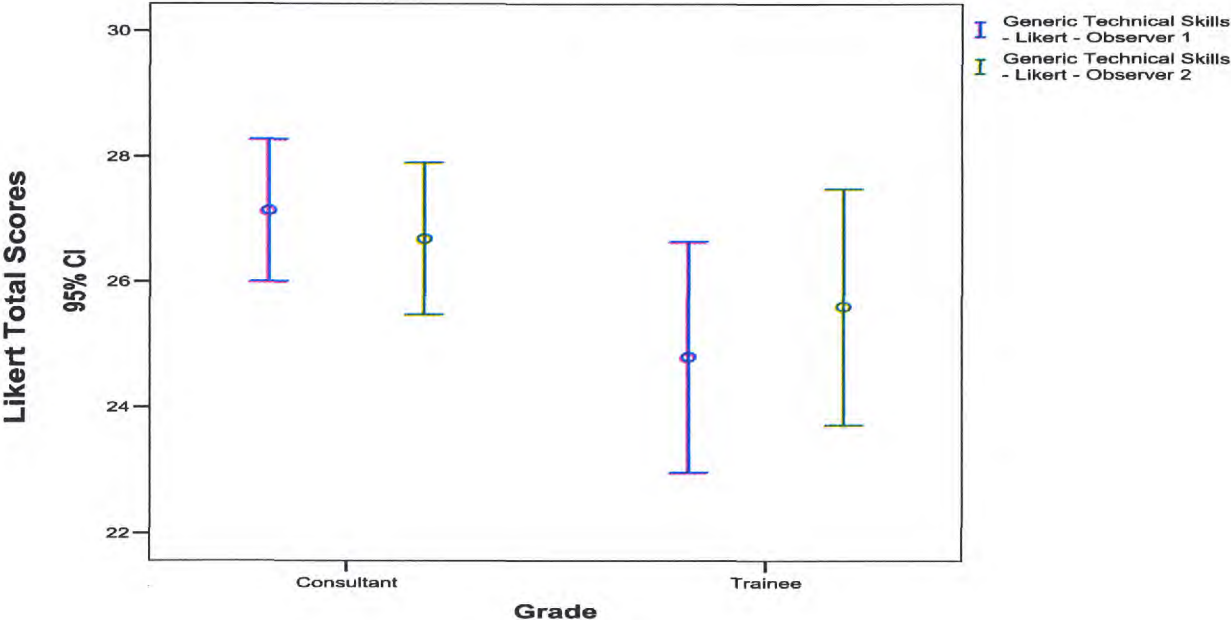
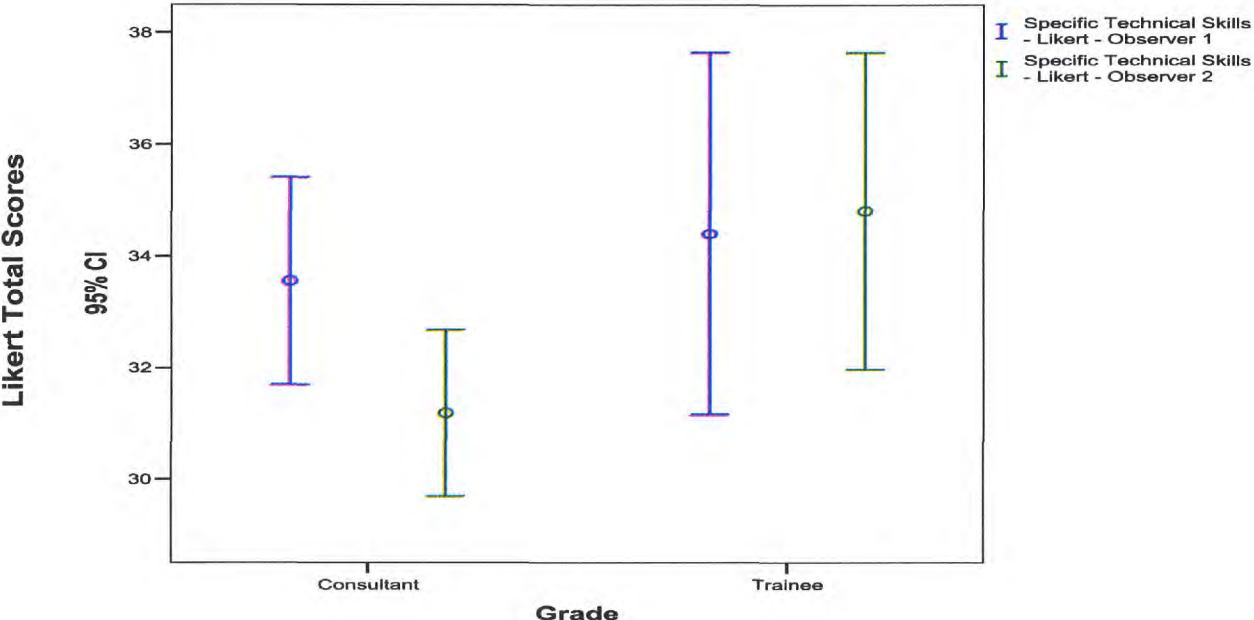


Figure 8.6 – Specific Likert Technical Skills Scoring of Consultants & Trainees – Live Open SF



Checklist Scoring – Between Consultants & Trainees – Live Open SF

Generic & Specific Technical Skills – Live Open SF

We are going to apply different techniques from the ones used in the likert scoring assessment for the reason mentioned in the reliability analysis, because the variables can not be considered continuous. We can test for differences between Consultants and Trainees through a non-parametric test for two independent samples i.e. Mann-Whitney using the average of the scores.

Lower scores relate to fewer errors in the generic technical skills checklist scale. In the minor, major and significant major events checklist (procedure specific technical skills), the **higher scores relate to more errors** being executed. Table 8.6 summaries the Mann-Whitney test results for the construct validity of each. The statistics are graphically displayed on Figures 8.7, 8.8 and 8.9.

Table 8.6 – Construct Validity of Checklists Technical Skills Scoring – Live Open SF

Technical Skill being Assessed	P Value	Conclusions
Generic Technical Skills	0.634	No significant difference between Trainees & Consultants
Specific Technical Skills Minor Events	0.931	No significant difference between Trainees & Consultants
Specific Skills Technical Major & Significant Major Events	0.071	No significant difference between Trainees & Consultants

Figure 8.7 – Generic Checklist Technical Skills Scoring of Consultants & Trainees – Live Open SF

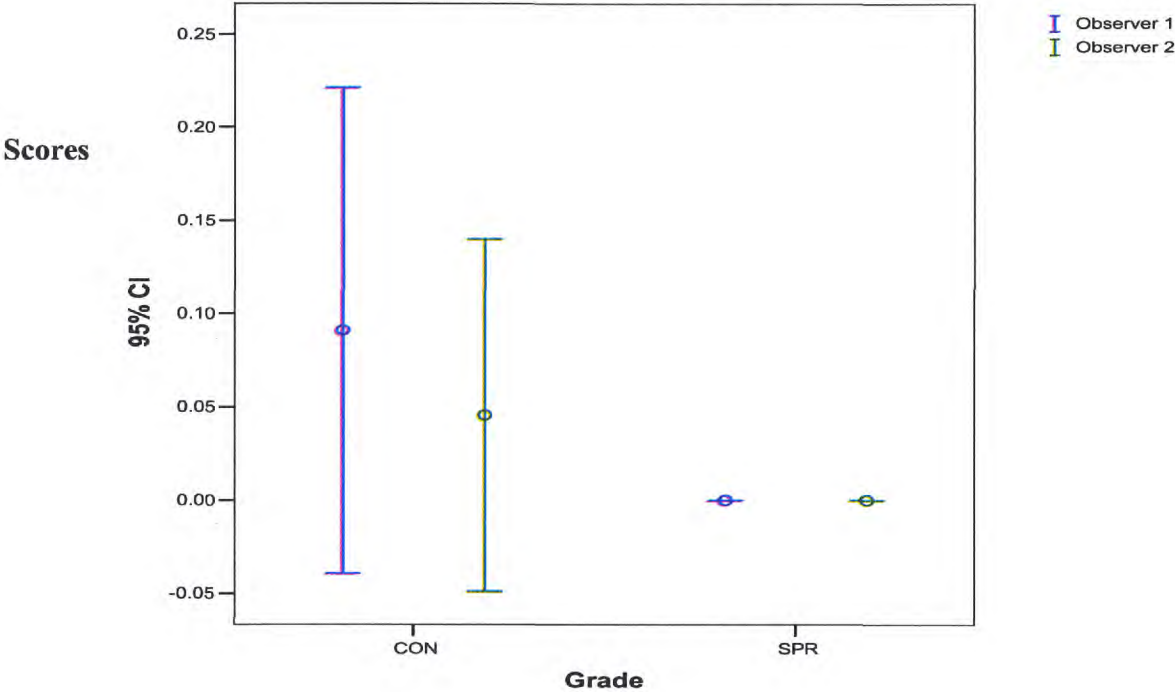


Figure 8.8 – Specific Minor Checklist Technical Skills Scoring of Consultants & Trainees – Live Open SF

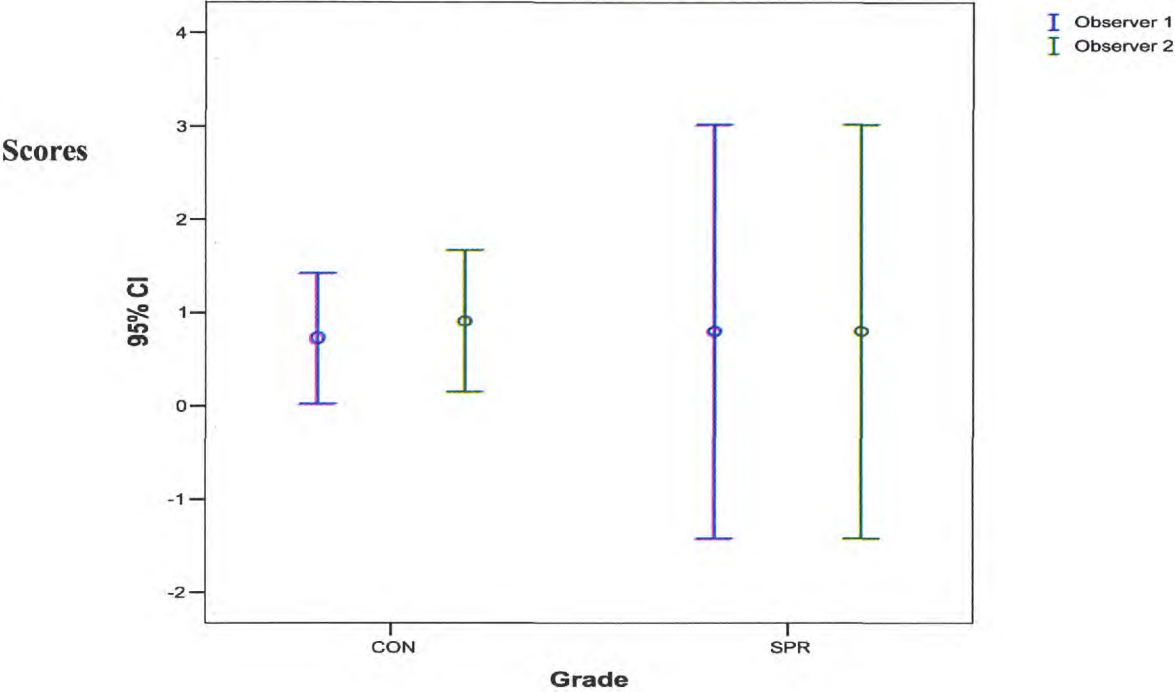
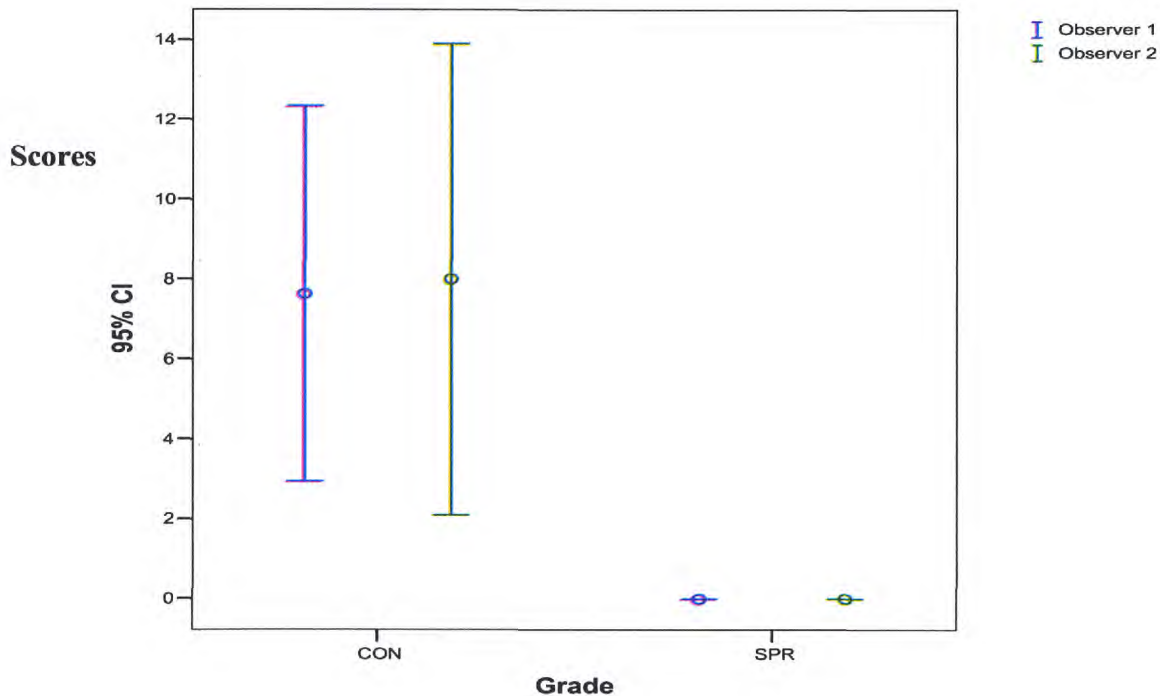


Figure 8.9 – Specific Major & Significant Major Checklist Technical Skills Scoring of Consultants & Trainees – Live Open SF



Reliability – Consultants & Trainees – Live & Simulated Open SF

We cannot assess reliability as there was only one observer who assessed the combination of live and synthetic open operations. We assume reliability is good for observer 1 (Specialist Registrar) on the basis of the good reliability between observers 1 and 2 in the live open sapheno-femoral junction ligation (first part of this Chapter 8), live open inguinal hernia repair (Chapter 7) and live laparoscopic cholecystectomy (Chapter 9) using the likert and checklist scoring. A second observer was not available to assess the simulated open SF operations.

Validity Between Consultants & Trainees – Live & Simulated Open SF

As there was small number of operations in the live open group with few trainees, it was decided to combine the live and simulated open sapheno-femoral junction ligations together. There were then 22 Consultant episodes and 25 Trainee episodes.

Likert Technical Skills

Generic & Specific Likert Technical Skills Validity - Live & Simulated Open SF

Using Repeated Measures Analysis of Variance (ANOVA) for construct validity (Figure 8.10 & 8.11) shows that there is a significant difference between Trainees and Consultants in generic technical skills, ($F = 5.491$, $p = 0.000$) but just no difference in specific technical skills, and ($F = 1.958$, $p = 0.056$). This means that on average Consultants perform generic technical skills better than Trainees, but on average Consultants (live operations) just perform specific technical skills the same as the Trainees (live & simulated operations).

Figure 8.10 - Generic Likert Technical Skills Scoring of Consultants & Trainees
Live & Simulated Open SF

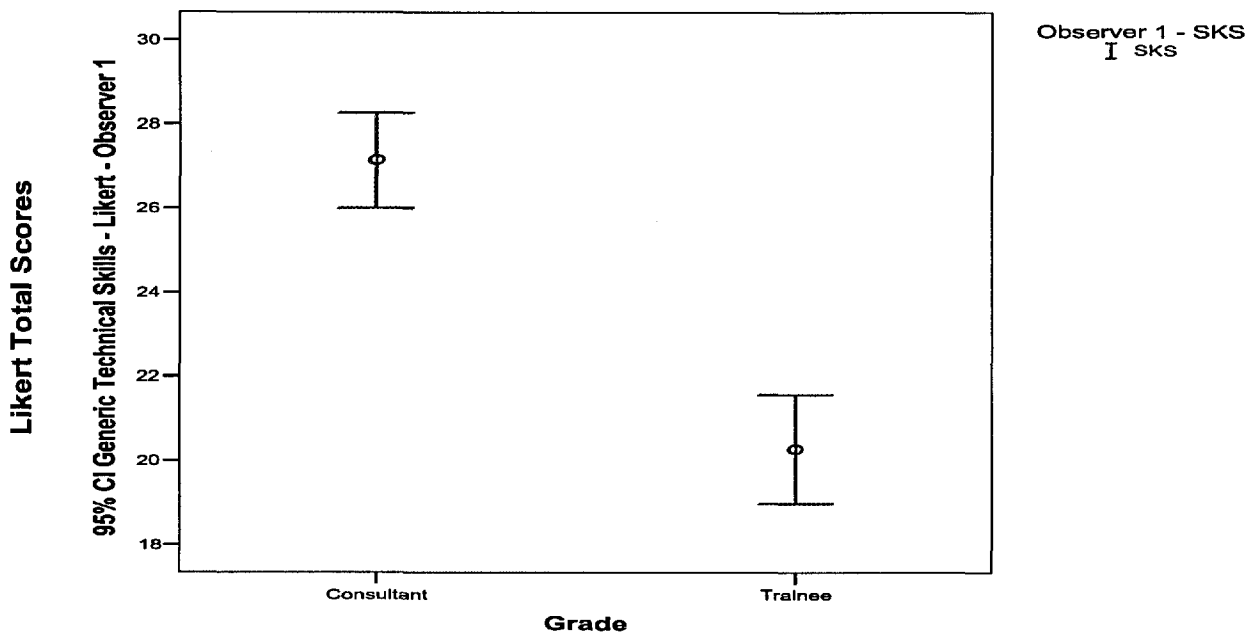
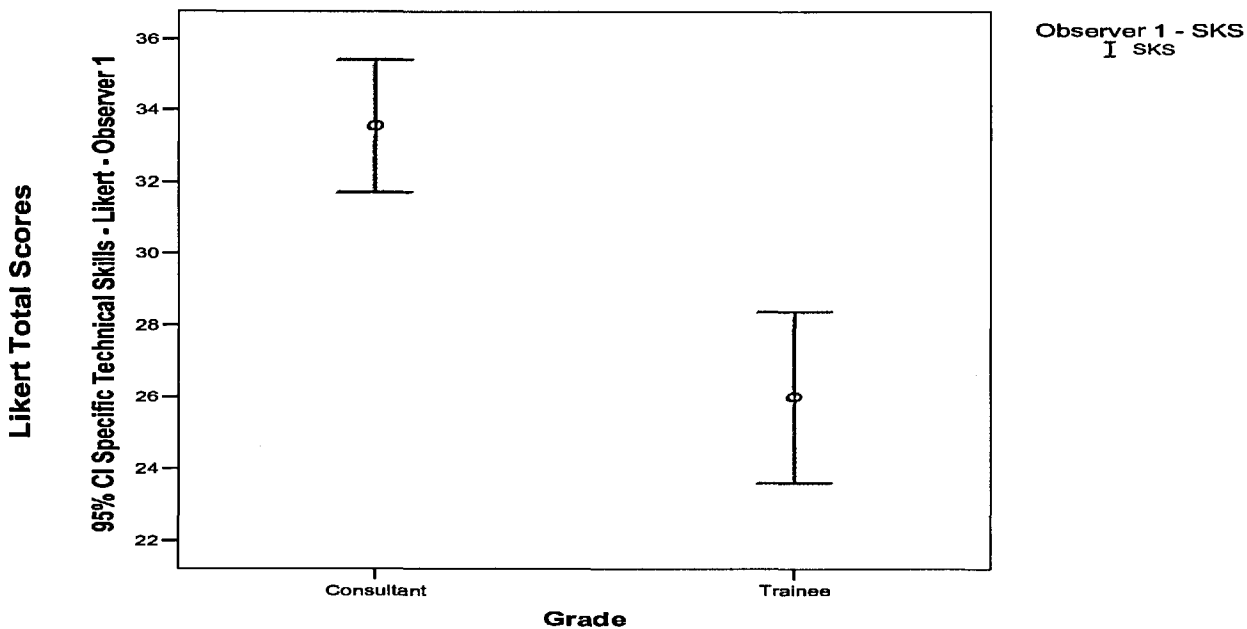


Figure 8.11 - Specific Likert Technical Skills Scoring of Consultants & Trainees
Live & Simulated Open SF



Checklist Technical Skills

Generic & Specific Checklist Technical Skills Validity - Live & Simulated Open SF

We are going to apply different techniques from the ones used in the likert scoring assessment for the reason mentioned in the reliability analysis, because the variables can not be considered continuous. We can test for differences between Consultants and Trainees through a non-parametric test for two independent samples i.e. Mann-Whitney using the scores (Table 8.7). The statistics are displayed in Figures 8.12, 8.13 and 8.14.

Table 8.7 – Validity of Technical Skills Checklists Between Consultants & Trainees in Live & Simulated Open SF Operations

Technical Skill being Assessed	P Value	Conclusions
Generic Technical Skills	0.344	No significant difference between Trainees & Consultants
Specific Technical Skills Minor Events	0.360	No significant difference between Trainees & Consultants
Specific Skills Technical Major & Significant Major Events	0.083	No difference between Trainees & Consultants

**Figure 8.12 – Generic Checklist Technical Skills Scores Live & Simulated Open SF
(Consultants & Trainees)**

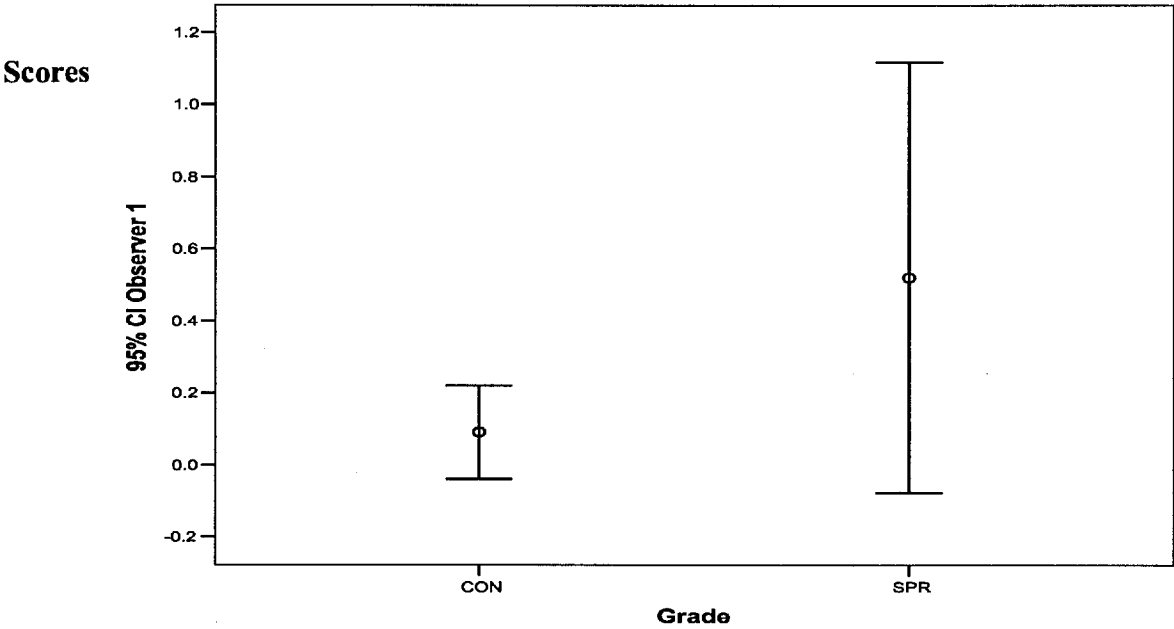


**Figure 8.13 - Specific Minor Technical Skills Checklist Scores Live & Simulated
Open SF (Consultants & Trainees)**

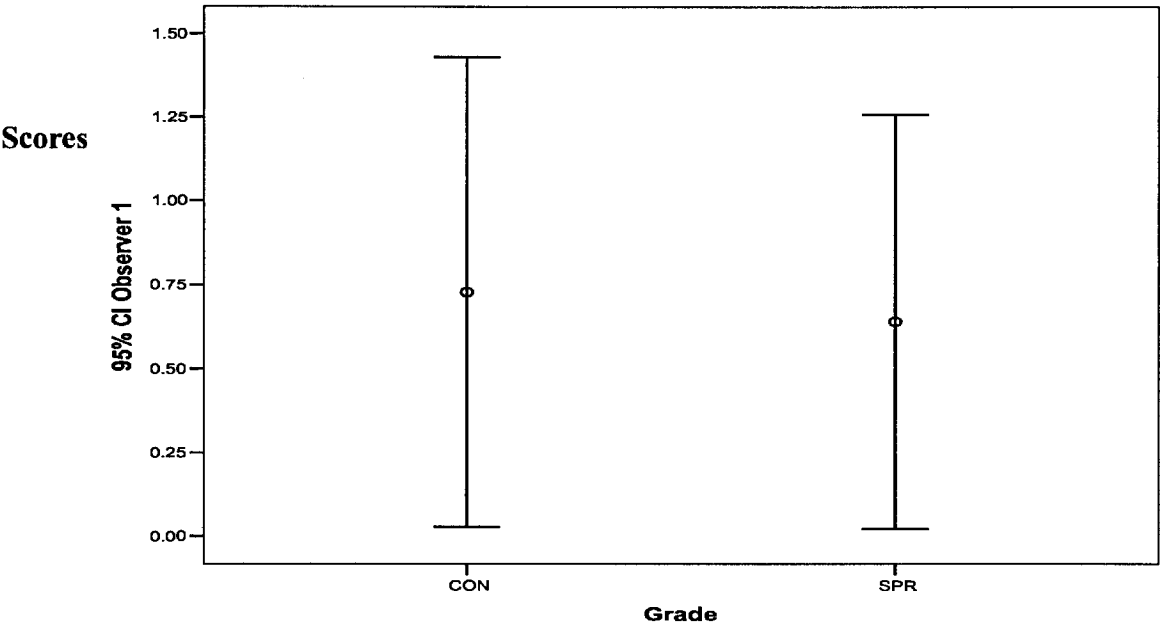
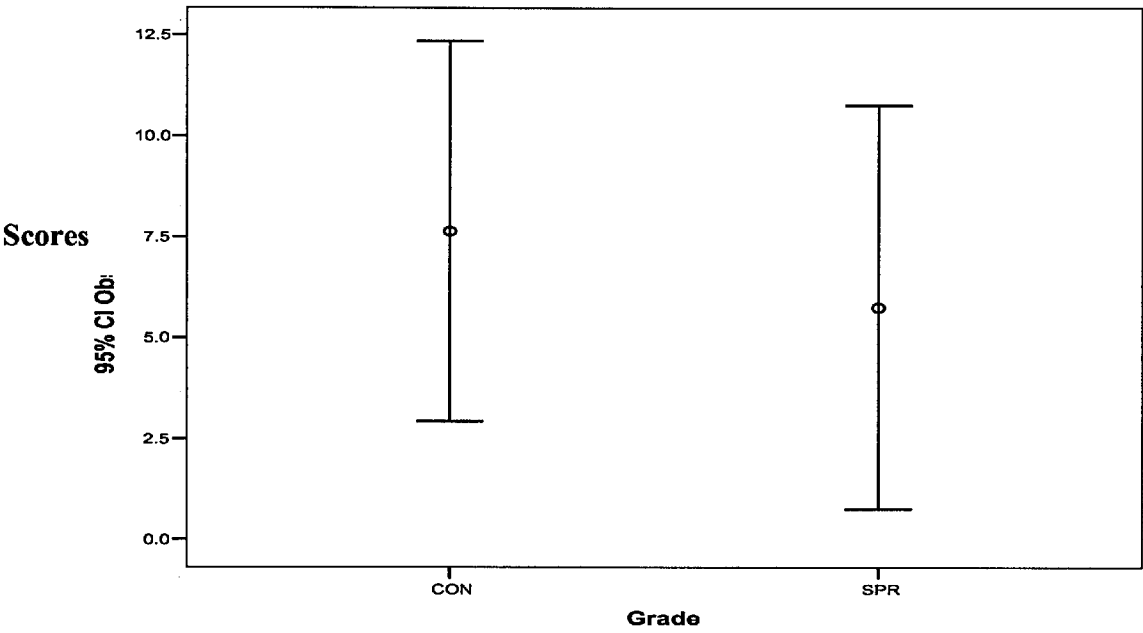


Figure 8.14 - Specific Major & Significant Major Technical Skills Scores Live & Simulated Open SF (Consultants & Trainees)



8.13 Conclusions

Reliability for the two observers was highly significant for the both likert and checklists scoring in the live operations. Due to the high reliability previously, one observer was used to assess the synthetic operations, as a second observer was not available.

Regarding live operative time and clinical grade between the Consultants and Trainees, there was no difference between the two groups. This is due to that fact that as a surgeon performs more operations his or her generic technical skills become more proficient and hence faster. As a result the process of executing each task and sub-task becomes significantly faster. Consultants had performed more than 150 primary sapheno-femoral junction ligations, and the Trainees were Senior and were approaching close to 150 operations.

Construct validity analysis for the likert scoring generic and specific assessment of technical skills in live operations were both not significant. As a group Consultants performed the operation the same compared to Senior Trainees regarding generic and specific technical skills using likert scoring. Construct validity analysis using Mann-Whitney for the checklist generic and specific minor technical skills were not significant respectively. Consultants as a group performed the operation the same compared to Trainees regarding generic and specific technical skills. The Senior Trainees were nearing Consultant status and had the same technical skill level as the Consultants. Therefore there was no difference in generic and specific technical skills using likert and checklist scoring.

As a result of a small number of Trainees in the live open group, a cohort of Trainees performing live simulated sapheno-femoral junction ligations on a validated bench model were added (Datta⁵ et al 2004).

Combining the two operations (live & synthetic) construct validity was significant in the likert generic scoring between Consultants and Trainees, but just not for specific likert scoring. Construct validity using the checklist for generic and specific technical skills were not significant between the Consultants and Trainees when the live and simulated open operations were combined.

This assessment tool of technical skills in open live and synthetic surgery seems to have face, content validities. However only concurrent, predictive and construct validities in the likert generic technical skills was demonstrated. The reason for this is that the live open study had significant unequal numbers in the Consultant and Trainee groups, and also the Trainees were senior and nearing Consultant status. The live simulated open operations only had Trainees and no Consultants. By combining the live and simulated open operations as a result did not produce the same results in the live open and laparoscopic studies, as we were not comparing the exact same operation. Also the numbers were small in all groups.

Table 8.1A – Generic Technical Skills Likert Scale for Sapheno-Femoral Junction Ligation

Category	1	2	3	4	5
Handling & Dissection of Tissues	Unnecessary force on tissues, causing frequent damage to tissues		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time & Motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument & Material Handling	Repeatedly makes tentative or awkward moves with instruments or material		Competent use of instruments/material although occasionally awkward		Fluid moves with instruments/material and no awkwardness
Use of Instruments & Materials	Frequently used inappropriate instruments & materials		Used appropriate instrument & materials for the task or subtask		Expertly used appropriate instruments & materials
Exposure of Technical Field	Consistently poor exposure		Good exposure most of the time		Expert exposure
Completion & Flow of Tasks & Subtasks	Frequently stopped operating or hesitant in next task or subtask		Forward planning with steady progression of tasks & subtasks		Effortless flow from one task to subtask

Table 8.1B – Specific Technical Skills Likert Scale for Sapheno-Femoral Junction Ligation

Specific Skill	2	4	6	8	10
Dissection of SFJ area	Did not expose area adequately with excessive trauma		Exposed area competently with minimal trauma		Exposed area expertly with no trauma
Tributaries	Could not or did not try to identify any tributaries		Identified all known tributaries. Did not seek other tributaries		Identified all known tributaries. Sought other possible tributaries
SFJ Clearance	Did not identify the SJ junction or used excessively traumatic dissection around femoral vein		Identified SF junction. Safely dissected tissues away from vessel. Reasonable clearance of vessel. Minimal trauma		Clearly identified SF junction. Expert dissection of tissues off the vessels. Atraumatic. Cleared well proximally and distally
SF Ligation	Did not ligate SF junction or ligated femoral vein or caused excessive encroachment onto femoral vein after SFJ ligation		Good knot tying whilst ligating the SJ. Minimal encroachment onto femoral vein following SFJ ligation		Excellent safe and secure ligation of the SFJ. Flush ligation with no encroachment onto femoral vein

Table 8.2A – Generic Technical Skills Checklists for Sapheno-Femoral Junction Ligation

Category	0	1
	Yes	No
Appropriate access to technical field		
Correct understanding of tissue planes		
Demonstrates sound knowledge of anatomy		
Confident & efficient progression of tasks & subtasks		
Uses instruments & materials efficiently & correctly		
Appropriate exposure/retraction of technical field		
Possesses good hand-eye co-ordination		
Handles tissues gently with minimal damage		
Correct suturing and knotting techniques		
Uses diathermy appropriately and safely		
Able to control bleeding efficiently & competently		
Appropriate wound management - closure & dressings		
Tasks or subtasks not rushed or too slow		

Table 8.2B – Specific Technical Skills Checklists for for Sapheno-Femoral Junction Ligation

(a) Procedure Specific Technical Skills – Checklist Scoring - Minor Events

	Yes (U) 6	Yes (C) 4	OK 0		Yes (U) 6	Yes (C) 4	OK 0
Incision not exactly over SFJ				Tributaries diathermied and still bleeding			
Tributaries avulsed							

(b) Procedure Specific Technical Skills - Checklist Scoring - Major Events

	Yes (U) 12	Yes (C) 8	OK 0		Yes (U) 12	Yes (C) 8	OK 0
Narrowing of femoral vein after SFJ ligated				Insecure tie on SFJ			

(c) Procedure Specific Technical Skills - Checklist Scoring – Significant Major Events

	Yes (U) 24	Yes (C) 16	OK 0		Yes (U) 24	Yes (C) 16	OK 0
Femoral vein injury				Less than 4 named tributaries tied			

(U) Yes Uncorrected

(C) Yes Corrected

OK = No error

Assessment Study

Chapter 9

Assessment of Technical Skills in Laparoscopic Cholecystectomy Surgery

9.10 Overview

As outlined assessing surgical technical skills in a structured manner is a topical issue. Laparoscopic surgery is relatively a new specialty, and has developed and expanded over the last decade. As new operations and techniques are developed in laparoscopic surgery it is important to assess how surgeons learn and perform these.

To date there has been no systematic attempt to comprehensively assess both generic and procedure specific technical skills in live laparoscopic surgery. In this present study we aim to develop and validate new assessment tools which can assesses both generic and procedure specific technical skills in laparoscopic live surgery.

9.11 Methods & Materials

Design of Study

This was a prospective unrandomised observational clinical study of operations, conducted between 2003 to 2004 in London, UK. Patients, Trainees and Consultants Surgeons were recruited at specific elective lists at the ACAD theatre at Central Middlesex Hospital and the main theatre at St Mary's Hospital.

Procedures of Study

Hierarchical Task Analysis of Laparoscopic Cholecystectomy

A description of a laparoscopic cholecystectomy is given in the previous hierarchical task analysis chapter in the PhD. Hierarchical task analysis was used to identify the key tasks and sub-tasks in laparoscopic cholecystectomy, and was the basis to construct the assessment tools in assessing specific technical skills. Hierarchical task analysis was constructed for laparoscopic cholecystectomy after referencing textbooks, articles, papers, web pages and expert panel discussions. This operation was chosen as it is performed by Trainees and Consultants in large numbers, one of the commonest general surgical laparoscopic operations and is a fundamental operation to be learnt and performed by both Trainees and Consultants.

Development of the Assessment Tool

Likert scales were then constructed individually for generic and procedure specific technical skills for the operation (Table 9.1A & 9.1B). Generic technical skills were scored incrementally by 1 point and procedure specific skills incrementally by 2 points. This was done after expert discussion, as the procedure specific technical skills were deemed more important to the generic technical skills as they can affect patient outcome and quality of the operation. Checklists (Table 9.2A & 9.2B) were also constructed for generic technical skills and procedure specific technical skills. In the later case, minor, major and significant major error scores were constructed and weighted after expert discussion. Generic skills errors were scored 0 point if correctly done and 1 point if incorrectly done. For minor errors, 0 points if no error was done, 4 points if done

incorrectly but then rectified and 6 points if incorrectly done and not rectified. For major errors, 0 points for no error being done, 8 points if incorrectly being done but rectified and 12 points if incorrectly being done and not rectified. For significant major errors 0 points for no error being done, 16 points if incorrectly being done but rectified and 24 points if incorrectly being done and not rectified. These scores reflected the severity of the specific technical skills error on the quality of the operation and possible patient outcome.

Data Collection

The video recordings of the operations were filmed using a Sony DCR-PC8E digital camcorder. Full length versions of operations were recorded onto digital video tape. These were then transferred without editing onto DVD, using Sony Click to DVD Software (Tokyo, Japan). Details of patients and subjects were recorded on computerized study forms, completed pre and postoperatively.

Assessing the Operations

Two surgeons with >12 years postgraduate surgical experience (Consultant Surgeon and Senior Specialist Registrar) assessed each operation blindly and independently. They were watched on large screen televisions and assessed using the likert and checklist technical skills assessment tool. All of the operations were seen, and only slow parts (non-essential) of the operations were fast forwarded.

Patients & Setting

All 100 patients were between ASA grades 1-3, BMI < 32, < 85 years old and were from either the ACAD centre, Park Royal, London W5 or St Mary's Hospital, London W2.

Subjects

The surgeons were divided into two groups, Consultants (performed > 150 laparoscopic cholecystectomy) and Registrar Trainees (performed < 150 laparoscopic cholecystectomy). They were all explained about the study and verbally consented to participate in the study, and had the option not to participate. The video camera operator did not give advice on technical aspects of the operation before, during or after the operation to the surgeons. All the Registrar Trainee operations were supervised.

Ethical Approval

Local ethics committee approval was achieved and each patient had the opportunity for written consent and the surgeons had verbal consent. Either the patient or the surgeon had the opportunity to decline to be included in the study.

Statistical Analysis

Data was collated in an Excel[®] database (Microsoft, Redmond, Washington, USA). Statistical analysis was carried out with SPSS[®] software statistical package (SPSS, Chicago, Illinois, USA). $P < 0.05$ was regarded as statistically significant.

For reliability analysis between the two observers two statistical methods were used. For likert scoring Intra-class Correlation Coefficient (ICC), also known as Cronbach α , and for checklists Kappa coefficients (k) were used. ICC and $k > 0.61$ was deemed statistically significant, with $p < 0.05$ (Appendix 7.3). Confidence intervals were set at 95%.

The data for validity analysis was parametric for likert scores so the Repeated Measures Analysis of Variance (ANOVA) was used. For Checklists it was non-parametric data and the Mann-Whitney test was used for comparison between the two groups of surgeons.

For the time taken to do the operation for the two groups of surgeons, the data was parametric. A Wilcoxon test was used to compare the two groups. A senior statistician from the Statistical Advice Department at Imperial College London analysed the data for reliability and validity with the primary researcher.

9.12 Results

Descriptive of Statistics & Assessments

There were 106 laparoscopic cholecystectomies recorded. Six had to be discarded due to video recording technical problems. 100 were then converted to DVD and analysed. A total of 18 surgeons participated in the study (Consultants = 10, Registrar Trainees = 8. 17 of whom were right-handed and one was left-handed. 17 were male surgeons and one was female). 5 Consultants were less than five years in their present grade and 5 Consultants were more than five years as a Consultant Surgeon. There were 2 registrar

trainees in years 5-6 and 6 trainees between years 1-4. There were 86 Consultants episodes, mean of 8 (range 25-1), 14 Trainees episodes, mean of 2 (range 4-1). Mean scores and ranges of the scores using likert and checklists of the two different grades of surgeons are shown in Table 9.3A & B.

Table 9.3A – Means & Ranges of Likert Scores of Surgeons – Lap Chole

Grade of Surgeon	Mean Likert Generic		Mean Likert Specific	
Consultants	24	Range 16-30	36	Range 23-49
Trainees	18	Range 14-25.5	30	Range 25-36

Table 9.3B – Means & Ranges of Checklist Scores of Surgeons– Lap Chole

Grade of Surgeon	Mean Checklist Generic		Mean Checklist Minor		Mean Checklist Major & Significant Major	
Consultants	0.40	Range 0-7	9.30	Range 0-24	4.05	Range 0-44
Trainees	1.50	Range 0-5	8.77	Range 4-16	4.31	Range 0-16

Regarding the patients there were 28 male and 72 female patients. Mean age was 52 (range 21-82 years) and mean BMI was 29.

There is an overlap between Trainees and Consultants in the mean likert generic and specific scores as well as the checklists for generic, minor and major technical skills. Comparing the means for the likert and checklists scores between the groups of surgeons,

it seems there is a larger numerical difference between the Consultants and Trainees using the likert than checklist scoring.

There were 33 generic technical skills errors performed by Consultants and 22 by Trainees. There were 88 minor errors which were corrected (82 by Consultants, 6 by Trainees) and there were 90 minor errors which were not corrected (75 by Consultants, 15 by Trainees). There was 43 major error which was corrected (36 by Consultants, 7 by Trainees), and there were 4 major errors which were not corrected (4 by Consultants, 0 by Trainees). There were no significant major errors done by either group of surgeons.

Full construct validity analysis will be done below, deciphering if there is a statistical difference between the Trainees and Consultants, and therefore examining the validation of the assessment tools.

Grading of Gallbladders & Technical Skills

4 laparoscopic cholecystectomies were converted to open operations, and 2 laparoscopic cholecystectomies had on-table cholangiograms. The gallbladders were graded by a standard clinical-pathological scale:

Grade 1 – Thin walled, no adhesions

Grade 2 – Thin walled with adhesions

Grade 3 – Thick walled

Grade 4 – Thick walled, chronically inflamed

Grade 5 – Thick walled, acutely inflamed

The grades of the gallbladders can affect the technical aspects of the operation. The mean operation grade and level of surgeon are depicted in Figure 9.1. There were 93 gallbladders graded 1-3 and 7 graded 4-5. Using non-parametric analysis, Chi-Square, there was no significant difference between the gallbladder grade and level of surgeon (Consultant or Trainee), $p > 0.05$. This was also the case for the Consultant subgroups (Figure 9.2), $p > 0.05$.

Figure 9.1 – Gallbladder Grade & Level of Surgeon

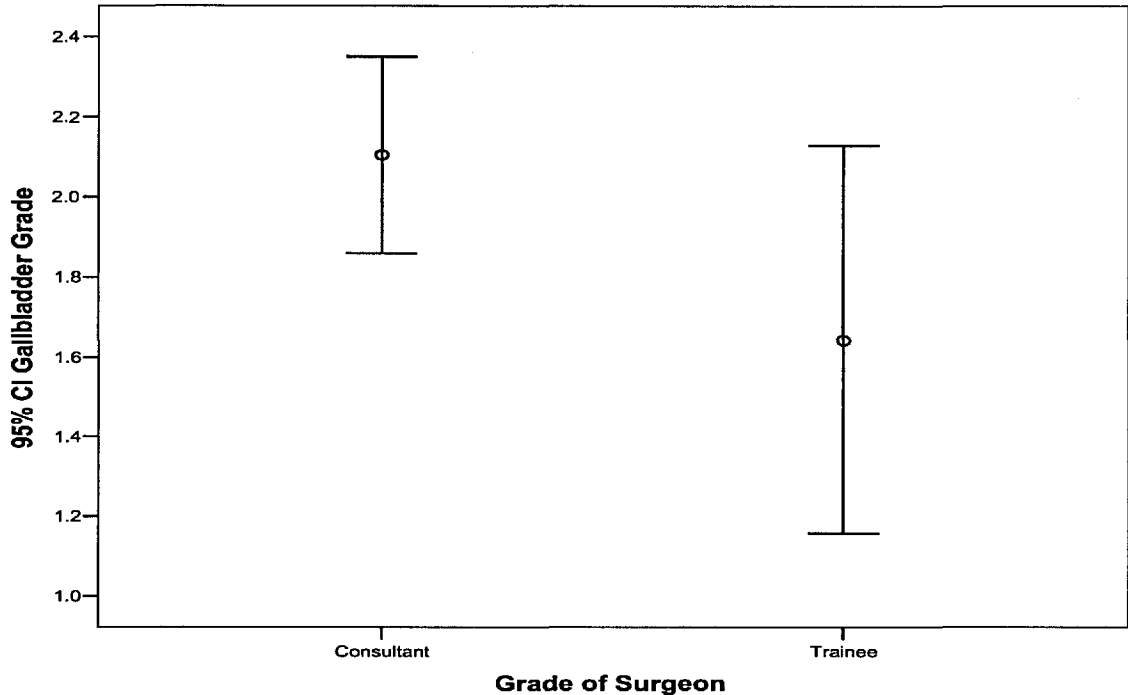
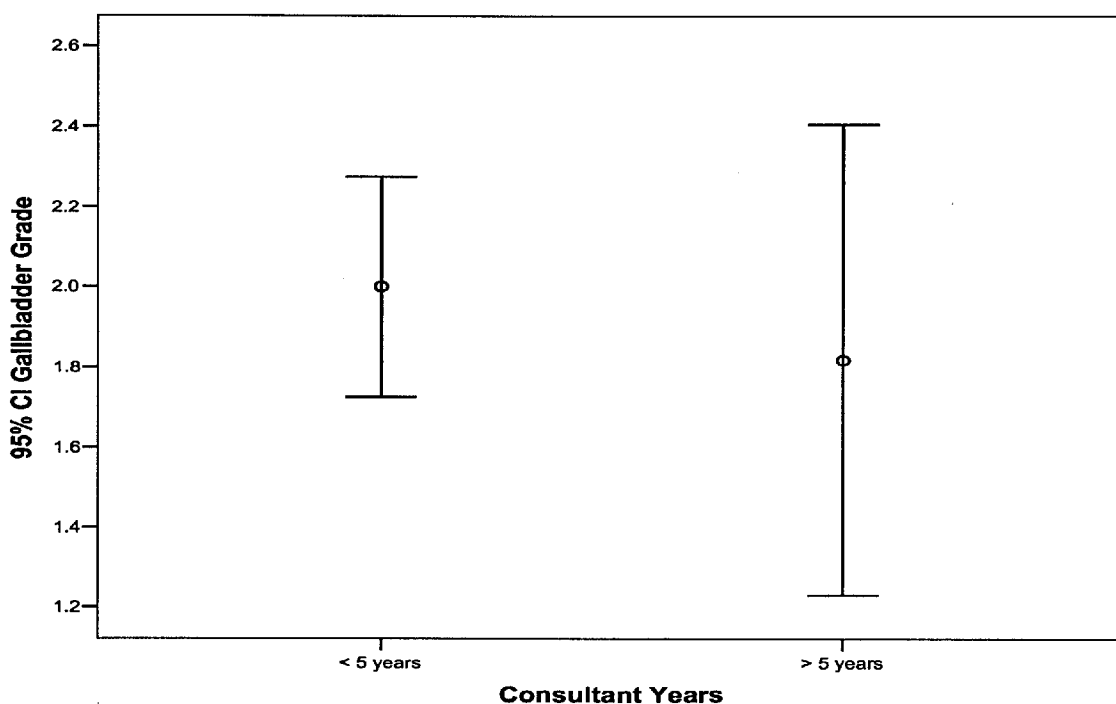


Figure 9.2 – Gallbladder Grade & Consultant Surgeon Years



Time Taken to Perform Operations

Mean time for the Consultant group to complete the operation was 37 minutes (range 15-120 minutes) and for the Trainee group was 50 minutes (range 20-80 minutes) Figure 9.3. Using parametric analysis, Wilcoxon test, the difference was significant, $p < 0.05$. Figure 9.4 depicts the Consultant Surgeon's operating time and gallbladder grade. The trainee's group did not have enough numbers for analysis of gallbladder grade. Using parametric analysis, Kruskal Wallis test, there was a significant difference between grade of the gallbladder and operating time of the Consultant Surgeons, $p < 0.05$. Using the Wilcoxon test there was a difference in operative time within the Consultant Surgeons group (Figure 9.5); ≤ 5 years mean 37 minutes (32-43 minutes), > 5 years 27 minutes (22-33 minutes), $p < 0.05$.

Figure 9.3 – Operating Time & Level of Surgeon

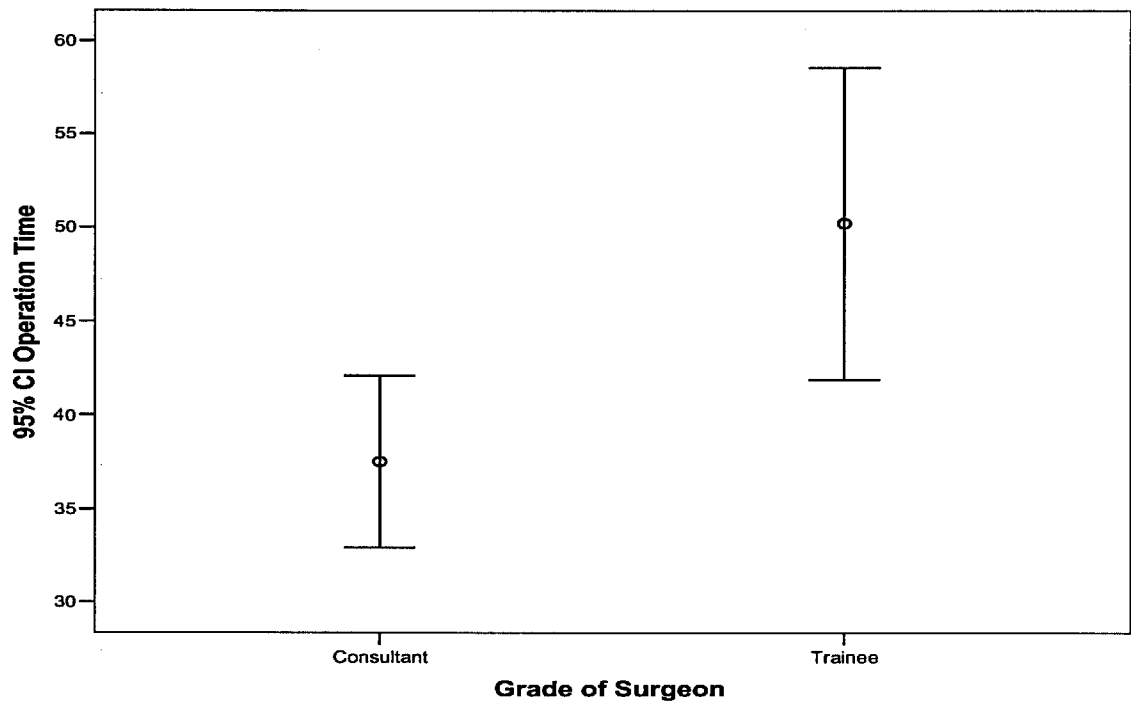


Figure 9.4 – Gallbladder Grade & Operating Time of Consultant Surgeons

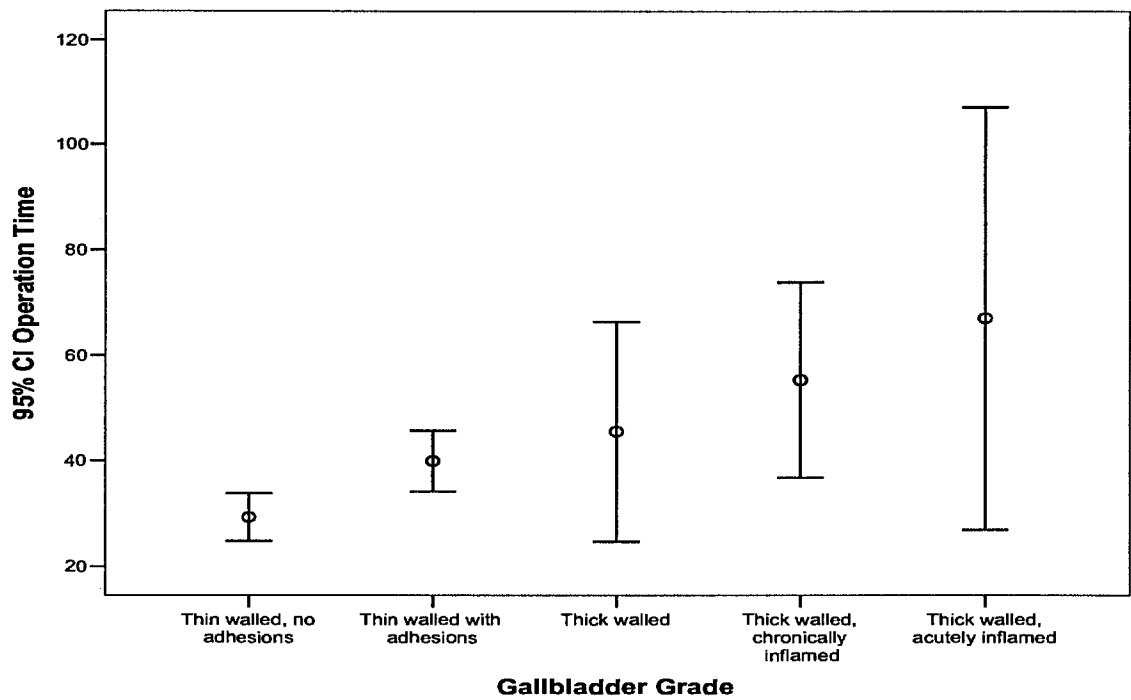
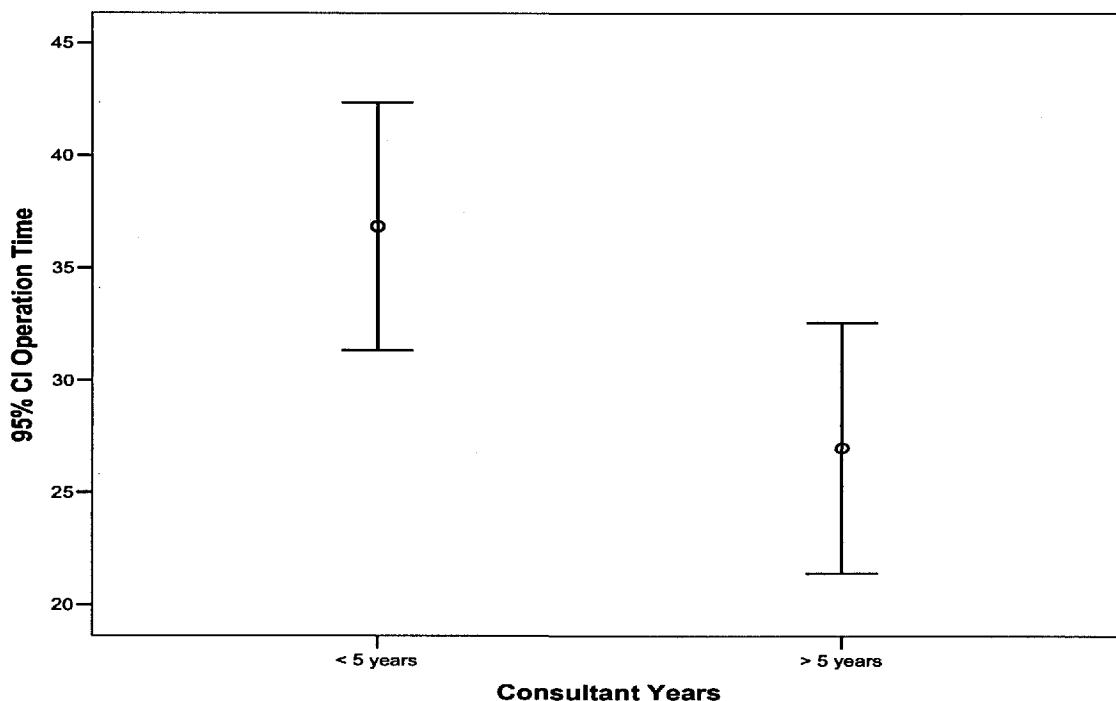


Figure 9.5 – Operating Time & Consultant Surgeon Years



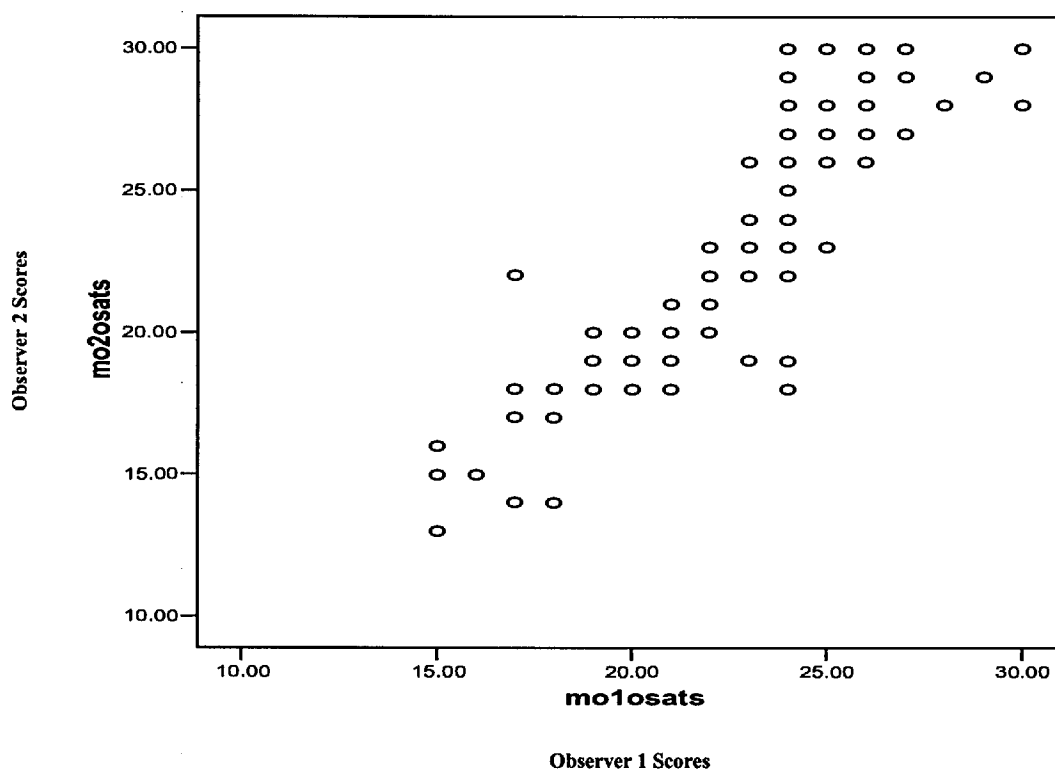
Likert Scoring Reliability – Between Observers

Generic Technical Skills – Likert Scoring

To assess reliability between the two observers for the likert technical assessment we use the intra-class correlation coefficient (ICC), also known as Cronbach α . Figure 9.6 depicts graphically the generic likert scores of observer 1 and 2. The Average Measure Intra-class Correlation ICC = 0.902, $p = 0.0000$. This means the reliability between the two observers is highly significant for generic likert scores.

Appendix 7.1 - for description

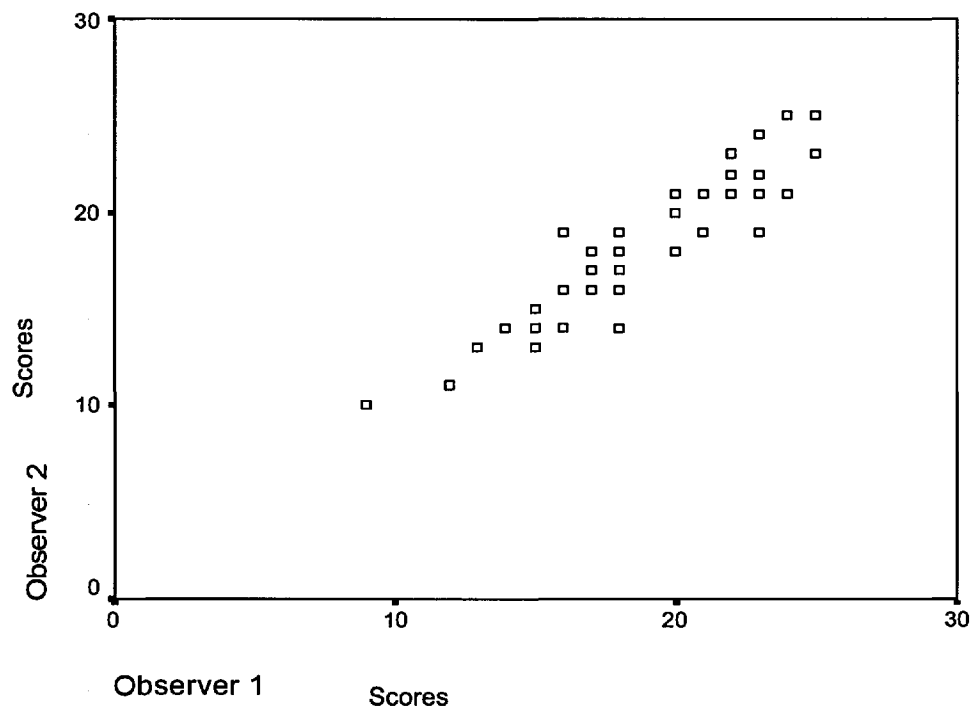
Figure 9.6 - Generic Technical Skills Likert Scoring Reliability Between Observers



Specific Technical Skills - Likert Scoring

The Average Measure Intra-class Correlation ICC = 0.919, $p = 0.0000$ for the two observers for the assessment of specific technical skills using likert scoring. Figure 9.7 depicts graphically the specific likert scores of observer 1 and 2. The statistics mean that the reliability between the two observers is highly significant.

Figure 9.7 - Specific Technical Skills Likert Scoring Reliability Between Observers



Checklist Scoring Reliability – Between Observers

To assess reliability between the two observers for the checklist technical assessment we use the weighted kappa coefficient (k). The reason why we use this method is explained in the following appendices. Table 9.4 summaries the kappa coefficients for all the checklists used. It can be seen that all the checklists have a high reliability between the two observers.

Appendix 7.2 – for description

Appendix 7.3 – for description

Table 9.4 – Reliability of Technical Skills Checklists Between Observers

Technical Skill being Assessed	Weighted Kappa Coefficients (k)	P Value	Conclusions
Generic Technical Skills	0.881	0.0000	Highly reliable between two observers
Specific Technical Skills Minor Events	0.924	0.0000	Highly reliable between two observers
Specific Skills Technical Major & Significant Major Events	0.859	0.0000	Highly reliable between two observers

Likert Scoring Technical Skills Validity – Between Consultants & Trainees

To assess validity for Trainees versus Consultants for the likert technical skills assessment we use Repeated Measures Analysis of Variance (ANOVA) using grade as a factor to test for differences between Trainees and Consultants.

Generic & Specific Technical Skills - Likert Rating

Figure 9.8 and 9.11 depicts graphically the statistical analysis between Trainees and Consultants. Using Repeated Measures Analysis of Variance (ANOVA) for construct validity shows that there is a significant difference between Trainees and Consultants in both generic (F-test = 32.587) and specific (F-test = 11.477) technical skills, $p < 0.05$. This means that on average Consultants perform generic and specific technical skills better than Trainees.

Figure 9.9 and 9.12 depicts the mean generic and specific technical skills scores using likert scoring for individual Consultant Surgeons.

To make more sense of this, the Consultant Surgeons were divided into two groups. Consultant Surgeons who had been appointed for 5 years or less, and those who were appointed for more than 5 years. There were 5 in each group. Figures 9.10 depicts the mean likert generic and Figures 9.13 specific technical skills scores for the two groups of Consultant Surgeons. Using Repeated Measures ANOVA for construct validity between the Consultant Surgeons groups, there was no difference for generic, $F = 1.043$ (p value = 0.443) and specific technical skills scores, $F = 1.386$ (p value = 0.177).

Figure 9.8 - Generic Technical Likert Scores of Consultants & Trainees

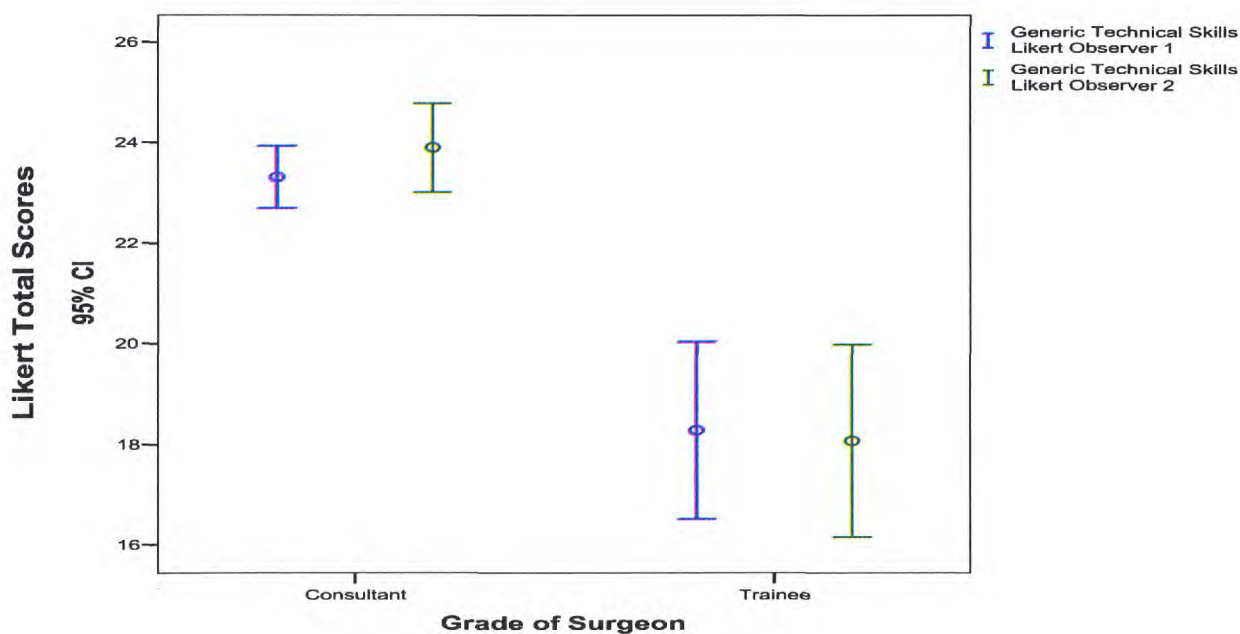


Figure 9.9 – Generic Technical Likert Scores & Individual Consultant Surgeons

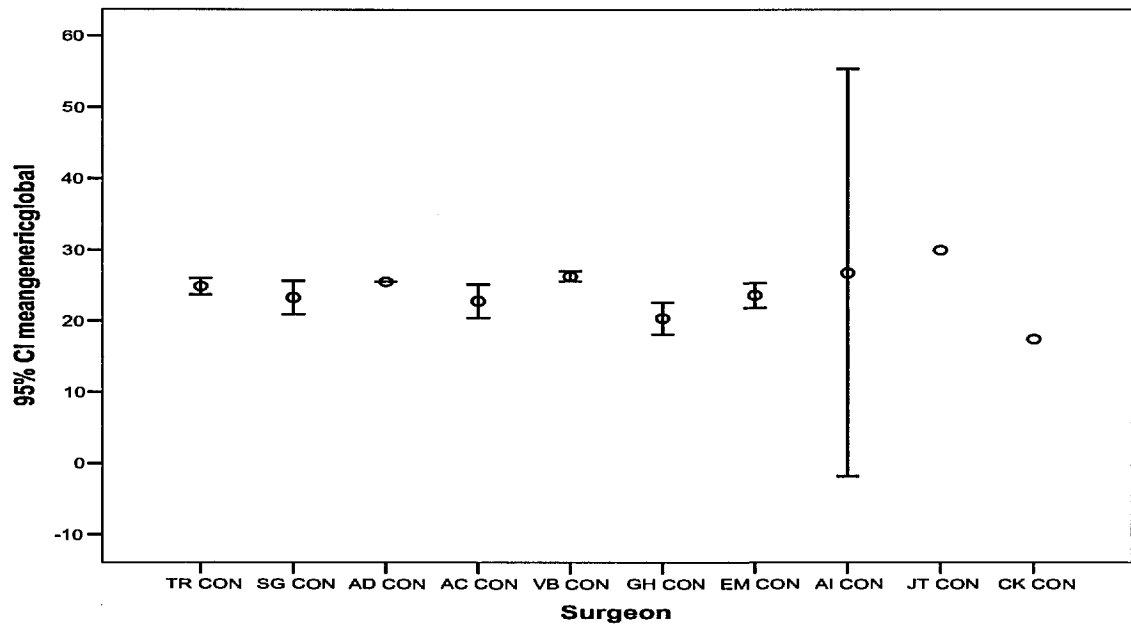


Figure 9.10 – Generic Technical Likert Scores & Consultant Surgeon Years

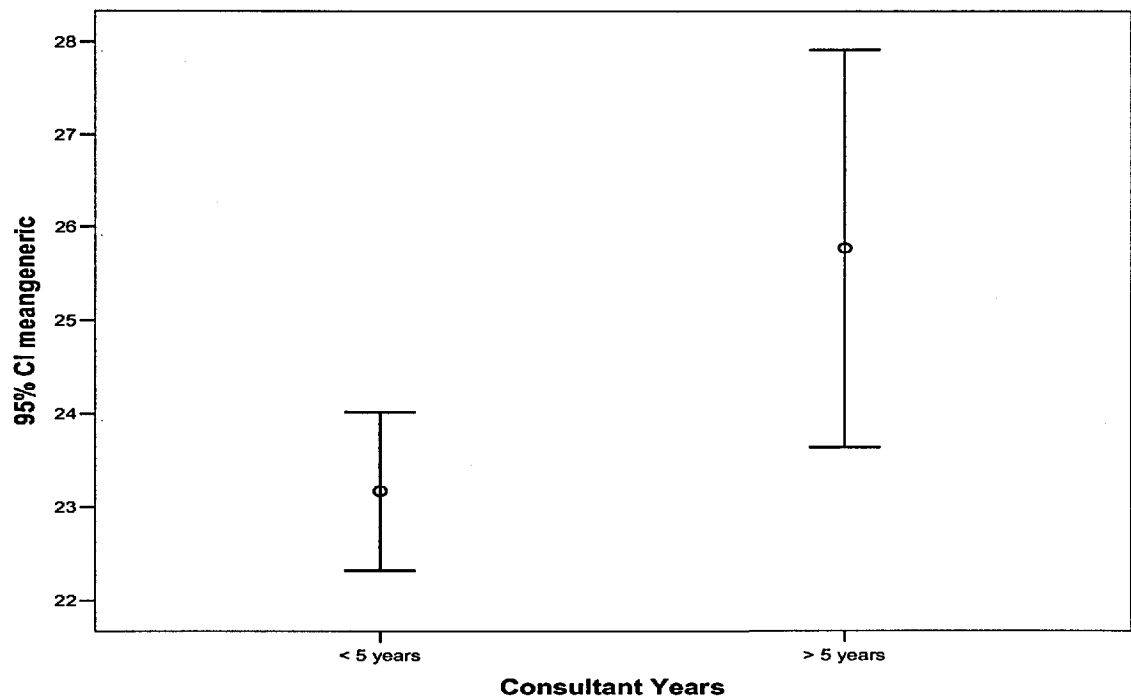


Figure 9.11 - Specific Technical Likert Scores of Consultants & Trainees

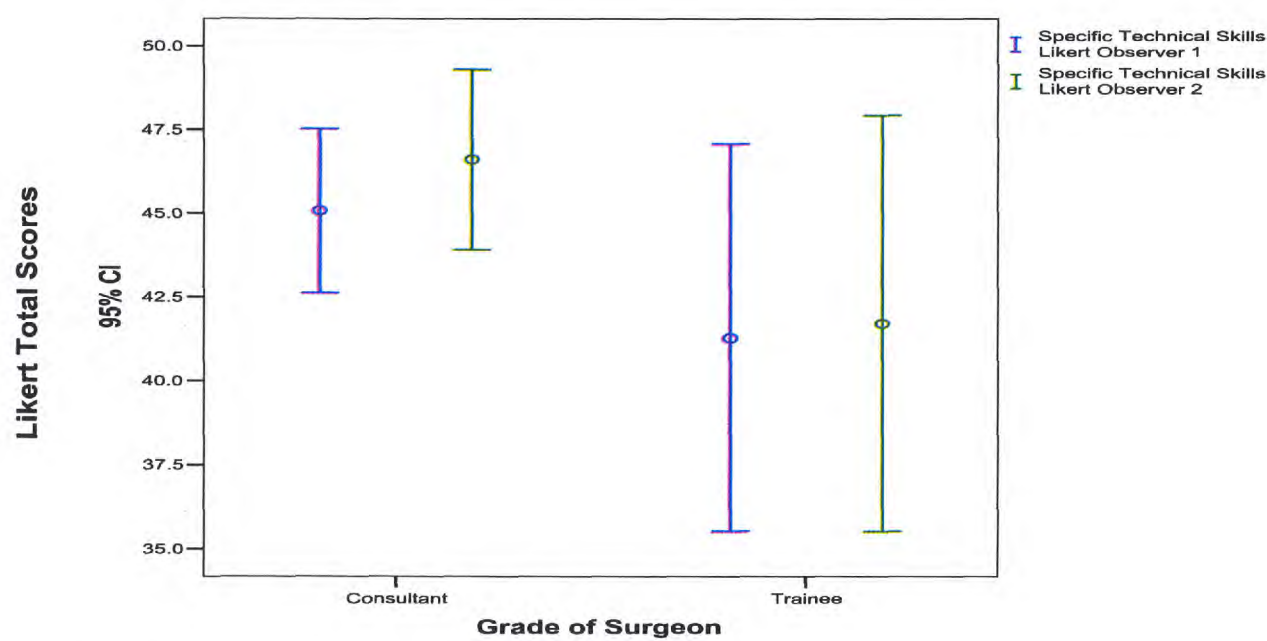


Figure 9.12 – Specific Technical Likert Scores & Individual Consultant Surgeons

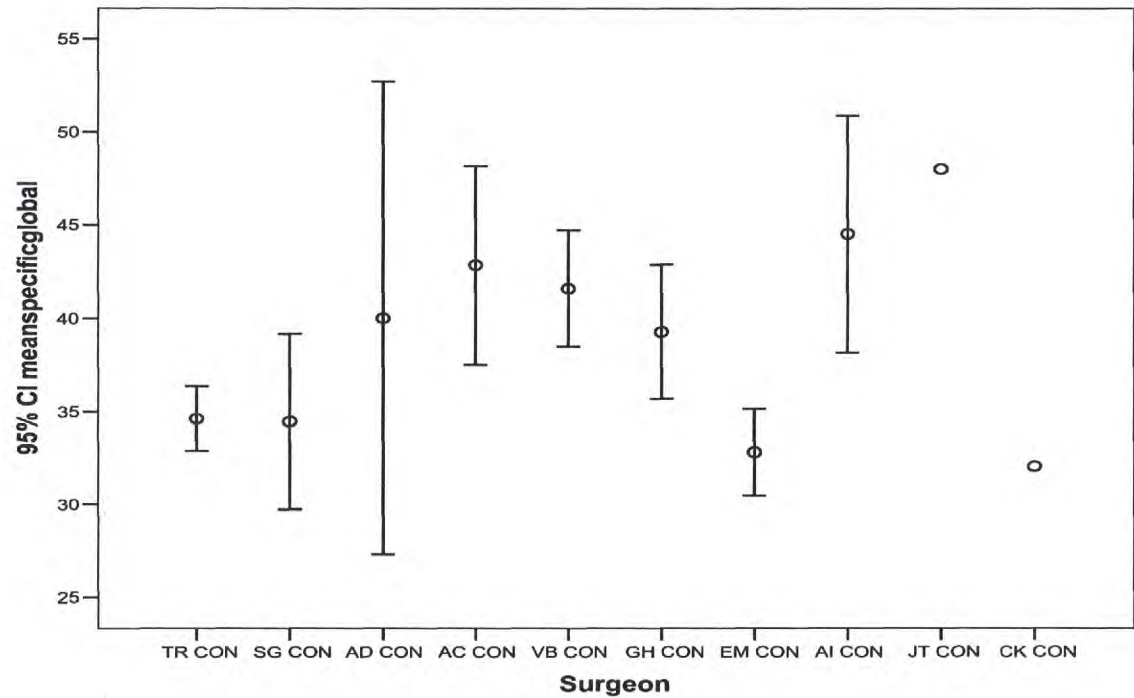
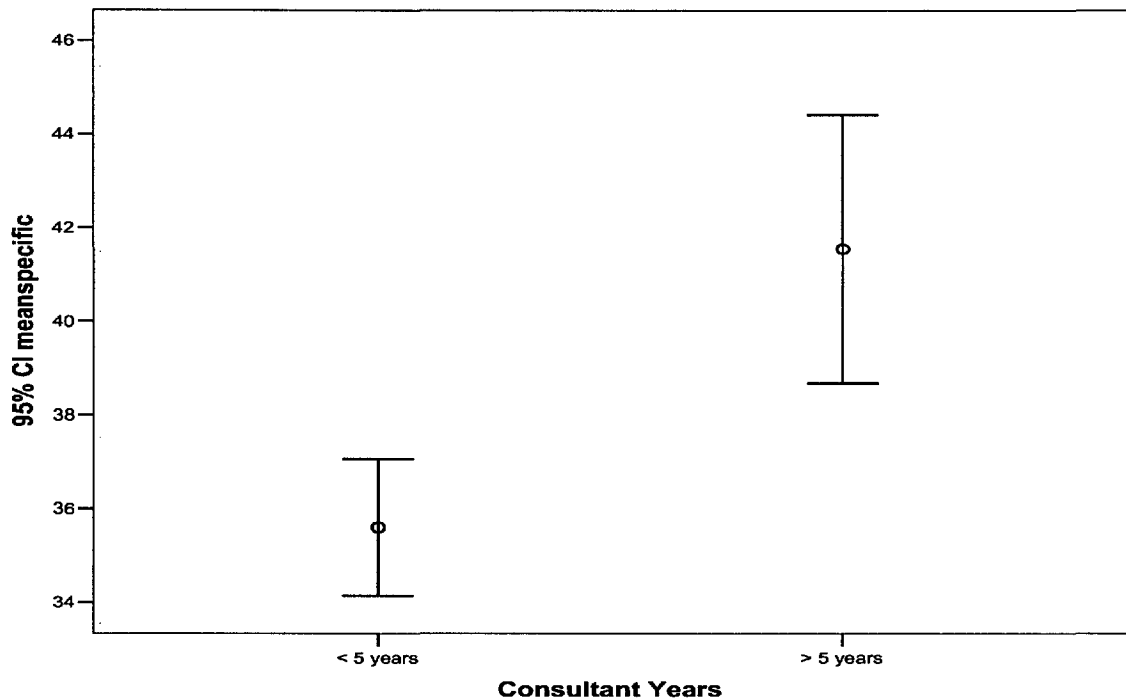


Figure 9.13 – Specific Technical Likert Scores & Consultant Surgeon Years



Checklist Scoring Technical Skills Validity – Between Consultants & Trainees

Generic & Specific Technical Skills

We are going to apply different techniques from the ones used in the likert scoring assessment for the reason mentioned in the reliability analysis, because the variables can not be considered continuous. We can test for differences between Consultants and Trainees through a non-parametric test for two independent samples i.e. Mann-Whitney using the average of the scores. It is useful to display a graph, such as error bar charts, with the central number being the mean checklist score with a 95% confidence interval (Figure 9.14 - 9.22). **Lower scores relate to fewer errors** in the generic technical skills checklist scale. In the minor, major and significant major events checklist (procedure

specific technical skills), the **higher scores relate to more errors** being executed. Table 9.5 and 9.6 summaries the Mann-Whitney test results for the construct validity of each.

Table 9.5 – Validity of Generic Technical Skills Checklists

Technical Skill being Assessed	P Value	Conclusions
Generic Technical Skills - Trainees & Consultants	0.0046	Significant difference between Trainees & Consultants
Generic Technical Skills - Consultant Surgeon Groups	0.511	No significant difference between Consultant Surgeon groups

Table 9.6 – Validity of Specific Technical Skills Checklists

Technical Skill being Assessed	P Value	Conclusions
Specific Technical Skills Minor Errors - Trainees & Consultants	0.273	No significant difference between Trainees & Consultants
Specific Technical Skills Minor Errors - Consultant Surgeon Groups	0.001	Significant difference between Consultant Surgeon groups
Specific Skills Technical Major & Significant Major Errors - Trainees & Consultants	0.602	No difference between Trainees & Consultants
Specific Skills Technical Major Errors - Consultant Surgeon Groups	0.160	No significant difference between Consultant Surgeon groups
Specific Skills Technical Significant Major Errors - Consultant Surgeon Groups	1.00	No significant difference between Consultant Surgeon groups

Figure 9.14 – Generic Technical Checklist Scores of Consultants & Trainees

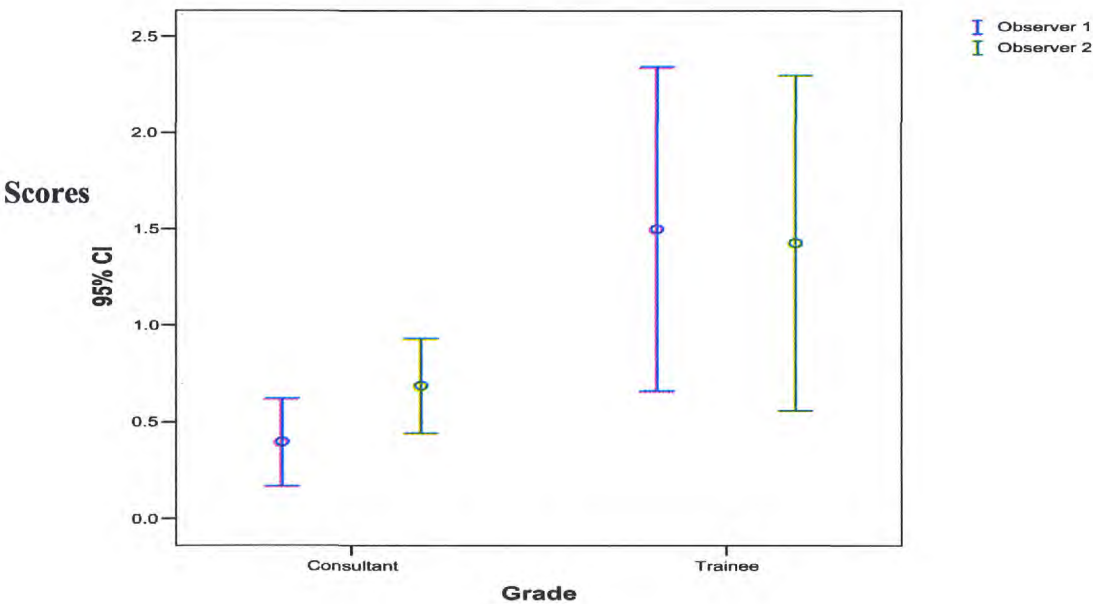


Figure 9.15 – Generic Technical Checklist Scores & Individual Consultant Surgeons

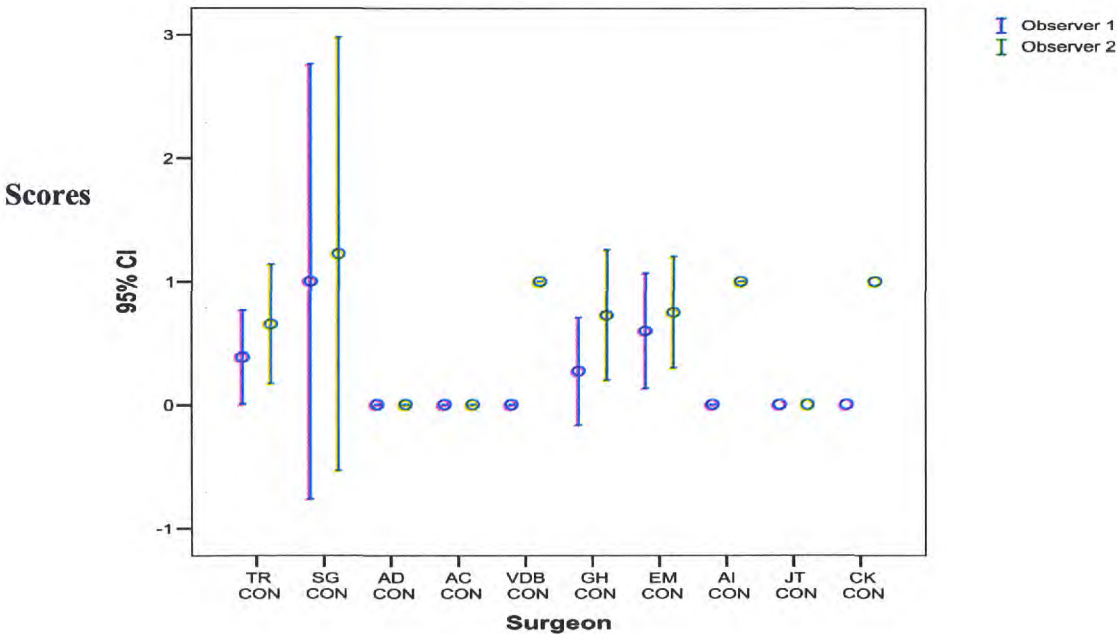


Figure 9.16 – Generic Technical Checklist Scores & Consultant Surgeon Years

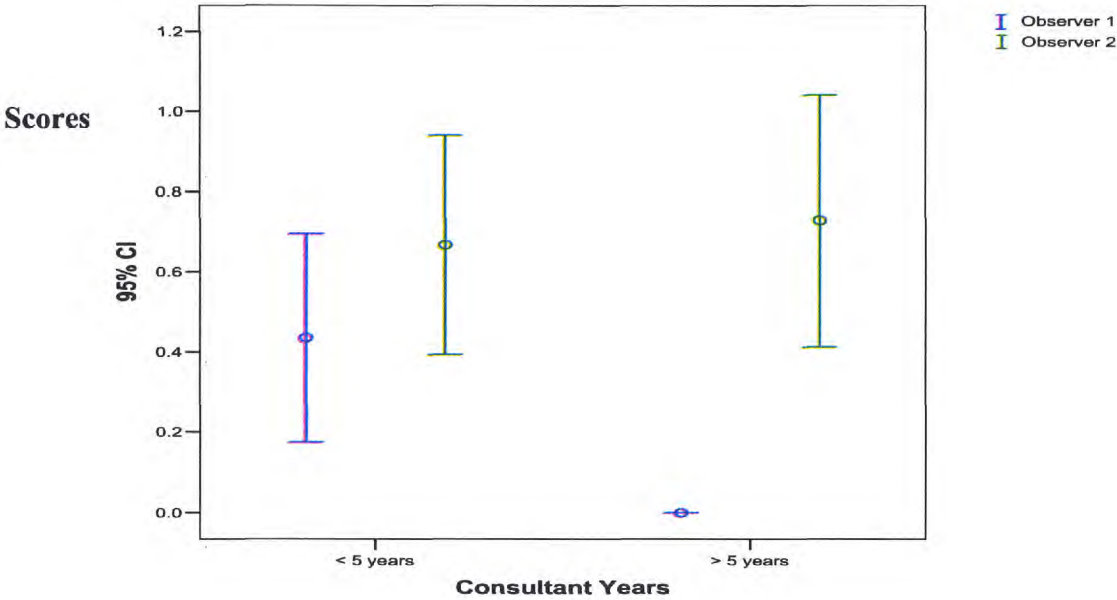


Figure 9.17 - Specific Minor Technical Skills Checklist Scores of Consultants & Trainees

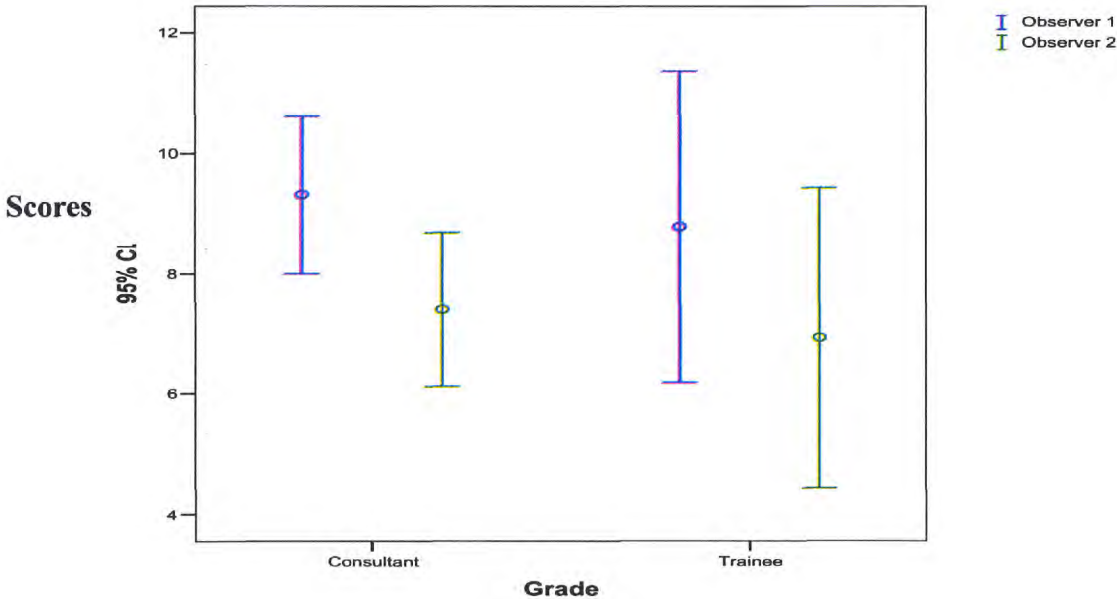


Figure 9.18 – Minor Technical Error Checklist Scores & Individual Consultant Surgeons

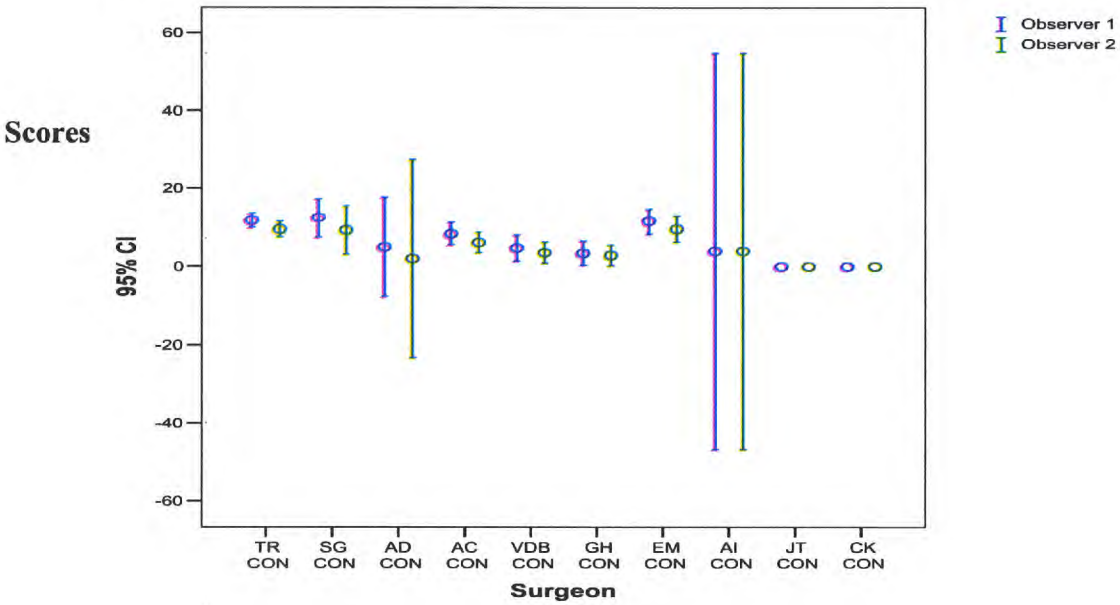


Figure 9.19 – Specific Minor Technical Skills Checklist Scores & Consultant Years

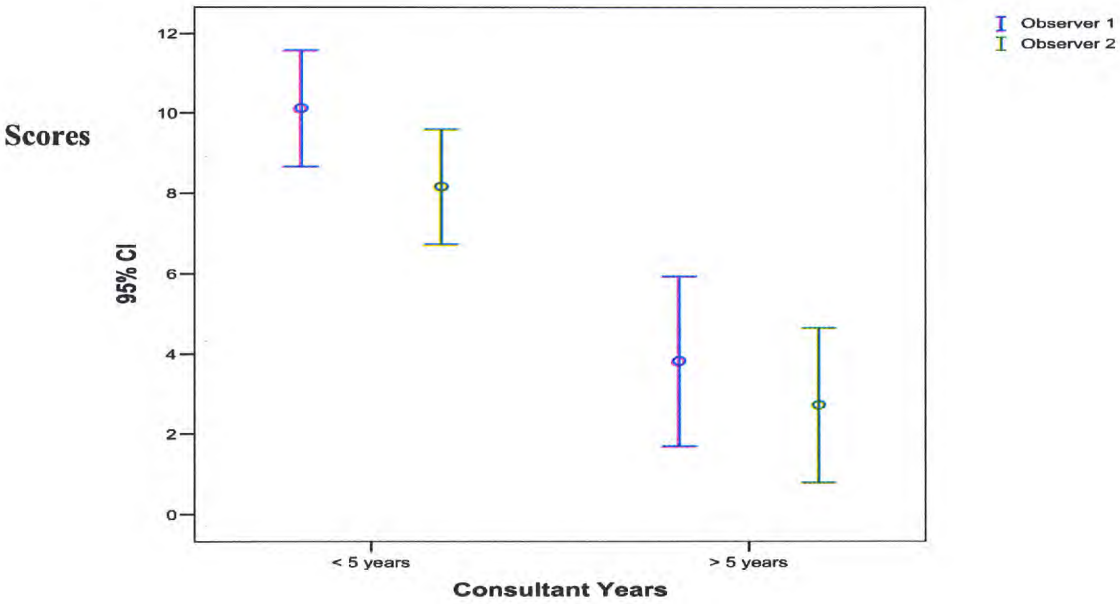


Figure 9.20 - Specific Technical Skills Checklist Major & Significant Major Scores of Consultants & Trainees

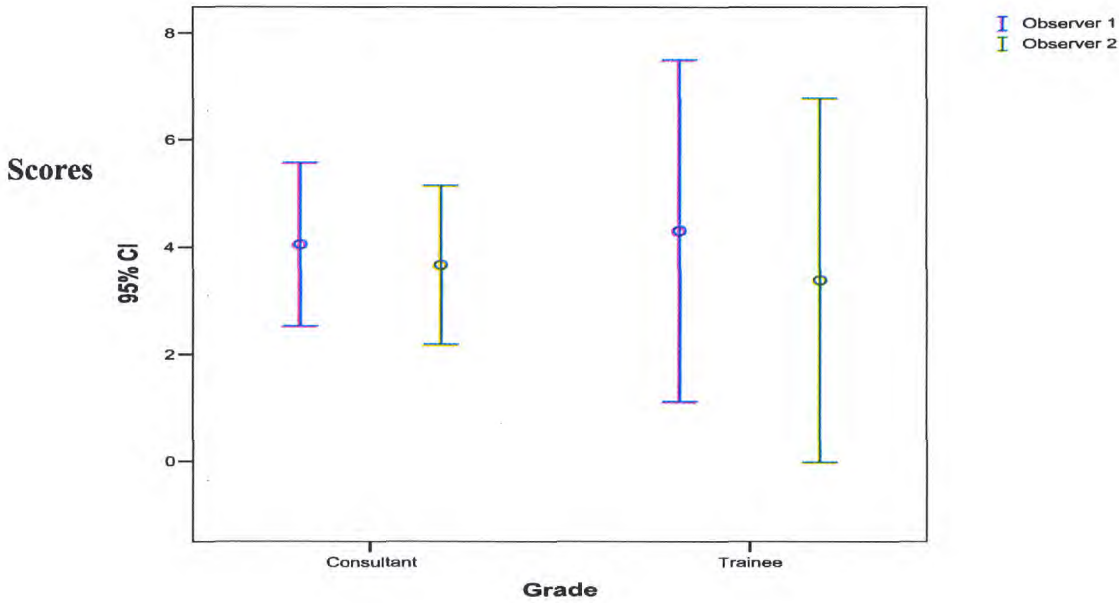


Figure 9.21 – Major Technical Checklist Scores & Individual Consultant Surgeons

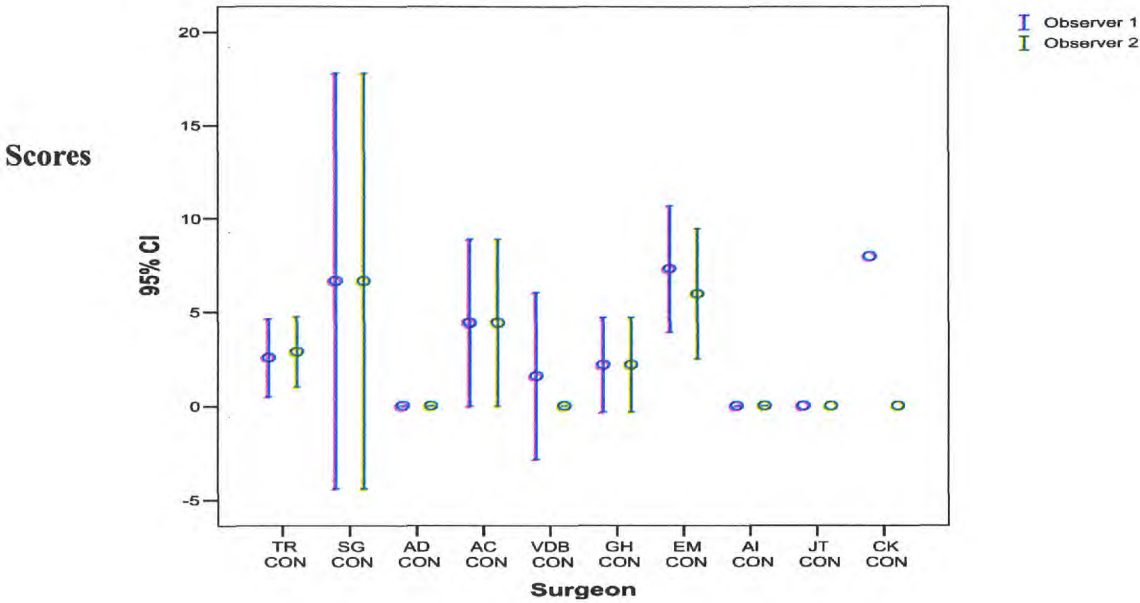
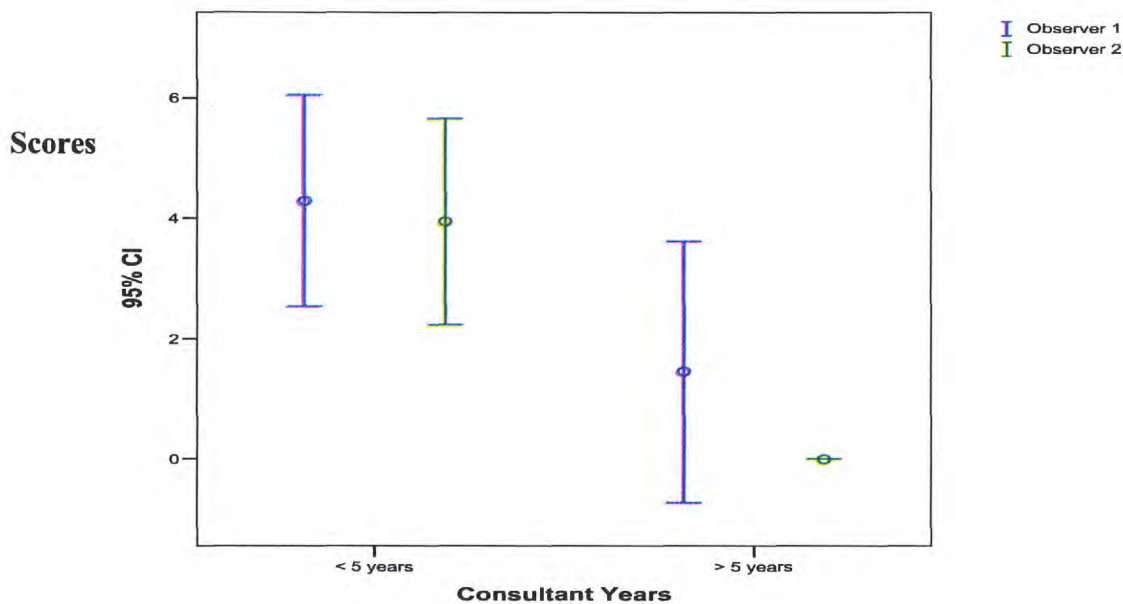


Figure 9.22 – Major Technical Skills Checklist Scores & Consultant Years



9.13 Conclusions

Reliability for the two observers was highly significant for the both likert and checklists scoring for the generic and specific technical aspects of the operation.

Construct validity analysis for the likert scoring generic and procedure specific assessment of technical skills were both significant. As a group, the Consultants performed the operation better compared to Trainees regarding generic and specific technical skills using likert scoring. Subgroup analysis did not show any difference between junior and senior Consultant Surgeons regarding likert scoring of generic and specific technical skills. Construct validity analysis for the checklist generic technical skills was significant. Consultants as a group performed the operation better compared to Trainees regarding generic technical skills using the checklist scoring. However for the

procedure specific technical skills (minor, major & significant major) checklist construct validity was not significant between these two groups. This later fact is probably related to the fact all Trainees were supervised.

Subgroup analysis of the junior and senior Consultant Surgeons did not show any difference except for minor technical errors. The junior Consultants were the same regarding generic technical skills and at major aspects of the operation compared to senior Consultants. However junior Consultants performed more minor technical errors and this is probably due to operative experience.

Comparing operating times the Consultants performed the operations faster than Trainees. This is due to that fact that as a surgeon performs more operations his or her generic technical skills become more proficient and hence faster. As a result the process of executing each task and sub-task becomes significantly faster. The Consultants had performed more than 150 laparoscopic cholecystectomies; the trainees had performed less than 150 operations. Subgroup analysis demonstrated that the Consultants took longer to perform operations with a higher clinical-pathological gallbladder grade. This is consistent as the higher grade of gallbladder, the technical aspects become more difficult and hence the surgeons took longer to perform the operation. Junior Consultants were slower than senior Consultants in performing the operation, and this probably due to the fact that the senior Consultants had operated more times than the junior Consultants. These assessment tools of technical skill in live laparoscopic surgery seems to have face, content, concurrent, predictive and construct validities in the likert scales and for the

checklists for only generic technical skills. It also demonstrates good reliability in the assessment of likert and checklists between the two observers. However comparing likert and checklist scoring it seems there is a larger numerical difference between Trainees and Consultants using the likert scoring than the checklist scoring.

All of the Consultant Surgeons in the study were deemed technically competent, hence they were not supervised. All of the Trainees were supervised, and this is a limitation of the study. There is no group difference appeared in the procedure specific technical skills checklist (major & significant major) categories. These categories would have an affect on patient outcome and quality of the final operation. A possible explanation for these results is that the supervised Trainees were prevented from making major errors as they were supervised by Consultants.

Table 9.1A – Generic Technical Skills Likert Scale for Laparoscopic Cholecystectomy

Category	1	2	3	4	5
Handling & Dissection of Tissues	Unnecessary force on tissues, causing frequent damage to tissues		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time & Motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument & Material Handling	Repeatedly makes tentative or awkward moves with instruments or material		Competent use of instruments/material although occasionally awkward		Fluid moves with instruments/material and no awkwardness
Use of Instruments & Materials	Frequently used inappropriate instruments & materials		Used appropriate instrument & materials for the task or subtask		Expertly used appropriate instruments & materials
Exposure of Technical Field	Consistently poor exposure		Good exposure most of the time		Expert exposure
Completion & Flow of Tasks & Subtasks	Frequently stopped operating or hesitant in next task or subtask		Forward planning with steady progression of tasks & subtasks		Effortless flow from one task to subtask

Table 9.1B – Specific Technical Skills Likert Scale for Laparoscopic Cholecystectomy

Specific Skill	2	4	6	8	10
Access & Port Insertion	Created clumsily & with difficulty		Created adequately		Created quickly & skilfully
Gall bladder retraction & exposure of Calot's triangle	Poor retraction & exposure of Calot's triangle		Satisfactory retraction & exposure of Calot's triangle		Expert retraction & exposure of Calot's triangle
Cystic Duct Dissection	Inadequate identification of cystic duct		Identified		Clearly identified
Cystic Duct Clipping & Transection	Clips not placed accurately. Bleeding		Clips placed proximally and distally adequately. Minimal bleeding		Clips placed expertly, proximally and distally, with adequate with length. No bleeding
Cystic Artery Dissection	Inadequate identification of cystic artery with clips not placed accurately.		Identified, clips proximally and distally adequate.		Clearly identified, clips proximally and distally adequate with enough length.
Gall bladder Fossa Dissection	Gall bladder perforated during procedure with bile and stone spillage. Excessive Liver bed bleeding		Gall bladder only minimally perforated during procedure with no bile and stone spillage. Liver bed dry with minimal bleeding.		Gall bladder entirely intact during whole procedure. Liver bed dry with no bleeding.
Extraction of gallbladder	Clumsily done with perforation of gallbladder		Adequately done with little trauma		Smoothly done with minimal trauma

Table 9.2A – Generic Technical Skills Checklists for Laparoscopic Cholecystectomy

Category	0	1
	Yes	No
Appropriate access to technical field		
Correct understanding of tissue planes		
Demonstrates sound knowledge of anatomy		
Confident & efficient progression of tasks & subtasks		
Uses instruments & materials efficiently & correctly		
Appropriate exposure/retraction of technical field		
Possesses good hand-eye co-ordination		
Handles tissues gently with minimal damage		
Correct suturing and knotting techniques		
Uses diathermy appropriately and safely		
Able to control bleeding efficiently & competently		
Appropriate wound management - closure & dressings		
Tasks or subtasks not rushed or too slow		

Table 9.2B – Specific Technical Skills Checklists for Laparoscopic Cholecystectomy

(a) Procedure Specific Technical Skills – Checklist Scoring - Minor Events

	Yes (U) 6	Yes (C) 4	OK 0		Yes (U) 6	Yes (C) 4	OK 0
Injury to gallbladder with bile spilled				Clip incompletely on cystic duct			
Liver injury, by diathermy				Misplaced clip fallen into abdomen			
Clip incompletely on cystic artery				Cystic artery or branches not identified initially			

(b) Procedure Specific Technical Skills - Checklist Scoring - Major Events

	Yes (U) 12	Yes (C) 8	OK 0		Yes (U) 12	Yes (C) 8	OK 0
Gallbladder injury with stones spilled				Unintentional cystic duct division			
Liver injury with bleeding				Cystic artery injury			
Loss of pneumoperitoneum due to surgery							

(c) Procedure Specific Technical Skills - Checklist Scoring – Significant Major Events

	Yes (U) 24	Yes (C) 16	OK 0		Yes (U) 24	Yes (C) 16	OK 0
Injury to other abdominal viscus				Duct Injury - CBD - Right Hepatic - Accessory			
Other major vascular injury							

(U) Yes Uncorrected

(C) Yes Corrected

OK = No error

Assessment Study

Chapter 10

Assessors of Technical Skills in Laparoscopic Cholecystectomy Surgery

10.10 Overview

The previous three chapters have demonstrated that assessing technical skills in open (live & simulated) and laparoscopic (live) surgery in a structured manner is feasible, and can be reliable and valid. In these studies we used two experienced surgeons with more than 12 years postgraduate surgical experience to be the assessors (observers) of the operations. They demonstrated good inter-rater reliability. However the grade and experience of the observer who can reliably assess the technical skills of an operation is yet to be determined.

It has been proposed that senior surgeons, retired surgeons or surgical trainees (assess their own or other trainees) could be used to assess the technical skills of Consultant Surgeons, non-training grade surgeons and surgical trainees. There are several possible contexts in which this might occur e.g. re-validation, research or clinical tools, self-development and self-appraisal of higher surgical trainees and Consultant Surgeons and the use of it in surgical technical skills teaching.

In this present study we will assess different types of assessors (observers) with varying degrees of surgical experience to assess an operation (laparoscopic cholecystectomy). We

have chosen to assess laparoscopic cholecystectomies which have already been assessed in Chapter 9, as they are most easily visualized on a video screen and therefore can be easily assessed.

This chapter is divided into two main individual studies.

1/ Study 1 - Training of Novices (BSc Students) to Assess Technical Skills in Laparoscopic Cholecystectomy

2/ Study 2 - Assessing Novice Surgeons (MSc Students) as Technical Skills Assessors in Laparoscopic Cholecystectomy

The main aim of the first study is to assess the learning and training effect of novices (no operative experience). This is to determine if novices can learn what are good or bad generic and specific technical skills when a surgeon performs a laparoscopic cholecystectomy operation. The second study's aim is to determine if novice surgeons are comparable to experienced surgeons in assessing technical skills in laparoscopic cholecystectomy.

10.1 Study 1

10.11 Methods

Design of Study

This was a prospective unrandomised observational study of a laparoscopic operation (laparoscopic cholecystectomy), conducted in 2003 to 2004 in London, UK at Imperial College Associated Hospitals. The technical skills assessors (Experienced Surgeons & BSc Students) were from Imperial College London.

Ethical Approval

Local ethics committee approval was achieved and each patient had the opportunity for written consent, the surgeons and assessors (Experienced Surgeons & BSc Students) had verbal consent. The patients, the surgeons or the assessors had the opportunity to decline to be included in the study.

Procedures of Study

Collection of Video Recordings

The video recordings of the operations were filmed using a Sony DCR-PC8E digital camcorder. Full length versions of operations were recorded onto digital video tape. These were then transferred without editing onto DVD, using Sony Click to DVD Software (Tokyo, Japan).

Assessing the Laparoscopic Operations by the BSc Students

The BSc students watched the DVD of the operation in a large lecture theatre on a large screen on a designated lecture day which was part of their course in December 2003. All of the operations were seen, and only slow parts (non-essential) of the operations were fast forwarded. They were not allowed to discuss between themselves of the scoring, trying to make the assessment as independent as possible. The identities of the surgeons were not given, so they assessed the DVDs blindly. The BSc students after watching the operation they were given instructions how to assess the generic and specific technical skills in laparoscopic cholecystectomy by two the experienced surgeons. They then watched the same operation again and re-assessed it. The mean scores from the BSc students of the technical assessments were then compared to the mean scores of the experienced surgeons.

Subjects

Fifteen 4th year medical students with limited surgical knowledge and experience were recruited into the study. Of the 15 BSc students there were 8 female and 7 male. They were intercalated BSc students at Imperial College London and had completed one year of clinical surgery prior to entering their BSc intercalated year. They had performed no laparoscopic operations, but had watched some laparoscopic cholecystectomies. They watched one operation performed by an expert surgeon and assessed the technical skills (before and after training) using the likert and checklist scoring for generic and specific technical skills as previously described in Chapter 9. They assessed one laparoscopic

cholecystectomy (Operations 1), which the experienced surgeons also assessed independently previously.

Training of BSc Students

As previously mentioned the BSc students assessed the operation initially without any instruction. Then after a break of 15 minutes they had training how to appropriately assess individual aspects of generic and specific likert and checklist scoring. This was demonstrated by showing the students video clips of corresponding assessment scores, so they had a visual benchmark to correspond to.

Gold Standard

The laparoscopic cholecystectomy which the BSc students assessed had previously been assessed by the two experienced surgeons. Since their inter-rater reliability in assessing the operation was high (Chapter 9), their scores were deemed to be the gold standard to which the BSc students' assessment scores were compared to.

10.12 Statistical Analysis

Data was collated in an Excel[®] database (Microsoft, Redmond, Washington, USA). Statistical analysis was carried out with SPSS[®] software statistical package (SPSS, Chicago, Illinois, USA). $P < 0.05$ was regarded as statistically significant.

Inter-rater reliability analysis between the BSc Students could not be done as there was a large amount of variance within the BSc student group and there were multiple assessors.

10.13 Results

Generic & Specific Ratings of Technical Skills

Table 10.1 summaries the mean scores of the BSc group (before and after training) collectively as well as those for the experienced surgeons for the relevant categories in assessing the technical skills of the operations. Full individual assessors' scores (before and after training) are given in the (Table 10.10 & 10.20). It can be seen that the BSc mean scores are closer to the experienced surgeons after training in 5 of the 6 categories, except for checklist generic. Also the range of the mean scores for each category after training reduces in variance for the BSc students.

Table 10.1 - BSc Students & Experienced Surgeons Mean Scores

Observer Group	Likert Generic Score & Range	Likert Specific Score & Range	Checklist Generic Score & Range	Checklist Specific Minor Score & Range	Checklist Specific Major Score & Range	Checklist Specific Significant Major Score & Range
BSc	13.1	35.7	11.2	17.1	21.0	24.0
Before Training	8-21	20-50	7-15	12-28	20-28	24-24
BSc	19.2	40.0	10.3	21.0	20.0	24.0
After Training	12-24	36-48	8-11	18-30	20-24	24-24
Experienced Surgeons	20.0	43.0	11.0	21.0	20.0	24.0
	20-20	42-44	11-11	20-22	20-20	24-24

**Table 10.10 – BSc Students (Before Training) & Experienced Surgeons
Technical Skills Assessment Scores**

Observer Type 1 = BSc 4 = Experienced Surgeons	Individual Assessor's Code	Likert Generic Score	Likert Specific Score	Checklist Generic Score	Checklist Specific Minor Score	Checklist Specific Major Score	Checklist Specific Significant Major Score
1.00	1.00	12.00	48.00	9.00	14.00	20.00	24.00
1.00	2.00	18.00	44.00	13.00	18.00	20.00	24.00
1.00	3.00	21.00	36.00	12.00	12.00	20.00	24.00
1.00	4.00	15.00	30.00	11.00	12.00	20.00	24.00
1.00	5.00	8.00	28.00	9.00	22.00	20.00	24.00
1.00	6.00	15.00	20.00	7.00	14.00	20.00	24.00
1.00	7.00	8.00	32.00	16.00	26.00	20.00	24.00
1.00	8.00	13.00	38.00	14.00	12.00	28.00	24.00
1.00	9.00	20.00	50.00	12.00	14.00	20.00	24.00
1.00	10.00	14.00	38.00	11.00	28.00	20.00	24.00
1.00	11.00	9.00	26.00	15.00	14.00	28.00	24.00
1.00	12.00	11.00	34.00	12.00	22.00	20.00	24.00
1.00	13.00	11.00	42.00	10.00	16.00	20.00	24.00
1.00	14.00	11.00	36.00	10.00	16.00	20.00	24.00
1.00	15.00	11.00	34.00	7.00	16.00	20.00	24.00
4.00	151.00	20.00	42.00	11.00	22.00	20.00	24.00
4.00	152.00	20.00	44.00	11.00	20.00	20.00	24.00

Table 10.20 – BSc Students (After Training) & Experienced Surgeons Technical Skills Assessment Scores

Observer Type 1 = BSc 4 = Experienced Surgeons	Individual Assessor's Code	Likert Generic Score	Likert Specific Score	Checklist Generic Score	Checklist Specific Minor Score	Checklist Specific Major Score	Checklist Specific Significant Major Score
1.00	1.00	23.00	44.00	10.00	18.00	20.00	24.00
1.00	2.00	19.00	45.00	11.00	18.00	20.00	24.00
1.00	3.00	21.00	36.00	9.00	22.00	20.00	24.00
1.00	4.00	12.00	38.00	11.00	22.00	20.00	24.00
1.00	5.00	20.00	41.00	9.00	20.00	20.00	24.00
1.00	6.00	20.00	36.00	11.00	18.00	20.00	24.00
1.00	7.00	17.00	41.00	11.00	28.00	20.00	24.00
1.00	8.00	17.00	44.00	10.00	22.00	20.00	24.00
1.00	9.00	24.00	44.00	11.00	18.00	20.00	24.00
1.00	10.00	16.00	34.00	11.00	30.00	20.00	24.00
1.00	11.00	21.00	30.00	11.00	22.00	20.00	24.00
1.00	12.00	21.00	36.00	8.00	18.00	20.00	24.00
1.00	13.00	21.00	48.00	10.00	20.00	20.00	24.00
1.00	14.00	17.00	38.00	11.00	20.00	20.00	24.00
1.00	15.00	19.00	44.00	10.00	20.00	24.00	24.00
4.00	151.00	20.00	42.00	11.00	22.00	20.00	24.00
4.00	152.00	20.00	44.00	11.00	20.00	20.00	24.00

Percentage of BSc Students Agreement to Experienced Surgeons

This was done to determine how accurate individual BSc Students were to the experienced surgeons in the scoring of technical skills. The mean scores of the experienced surgeons were compared with 95% accuracy (standard deviation +/- 5%) with those of individual BSc students before and after training (Table 10.2). Before training the accuracy is low in 4 of the 6 categories for the BSc students, major and significant major checklists excluded. After the training the accuracy improves in all 6 categories. The BSc students individually seem to improve after training.

Table 10.2 – Percentage Agreement of BSc Students to Experienced Surgeons

Operation Number	BSc Group	Likert Generic Score	Likert Specific Score	Checklist Generic Score	Checklist Specific Minor Score	Checklist Specific Major Score	Checklist Specific Significant Major Score
1	Before Training	13%	13%	47%	13%	87%	100%
1	After Training	53%	40%	80%	53%	93%	100%

In summary the BSc students improve in their accuracy in assessing technical skills after receiving training than without training. It is therefore possible to train novices to assess technical skills, however this falls short to acceptable levels of accuracy.

10.2 Study 2

10.14 Methods

Design of Study

See previous description.

Ethical Approval

Local ethics committee approval was achieved and each patient had the opportunity for written consent, the surgeons and assessors had verbal consent. The patients, the surgeons or the assessors (experienced surgeons & MSc students) had the opportunity to decline to be included in the study.

Procedures of Study

Collection of Video Recordings

See previous description.

Assessing the Laparoscopic Operations

The MSc students watched the DVD of the operation in a large lecture theatre on a large screen on a designated lecture day which was part of their course in March 2004. All of the operations were seen, and only slow parts (non-essential) of the operations were fast forwarded. They were not allowed to discuss between themselves of the scoring, trying to make the assessment as independent as possible. The identities of the surgeons were not given, so they assessed the DVDs blindly. The MSc students did not receive any training in how to assess the technical skills. The mean scores from the MSc students of the technical assessments were then compared to the mean scores of the experienced

surgeons. The experienced surgeons had previously watched and assessed the same operation as the MSc students.

Subjects

These were 18 junior surgeons (pre-MRCS) still in basic surgical training who were doing a part-time two year MSc in Surgical Science & Technology at Imperial College London. Of the 18 MSc students, there were 4 female and 14 male. They had not completed their MRCS (Member of the Royal College of Surgeons) examination which signifies the completion of Basic Surgical Training. They as a group had performed between 0-10 laparoscopic operations with supervision. They assessed 1 laparoscopic cholecystectomy (Operation 3), which the experienced surgeons also assessed independently previously.

Gold Standard

See previous description.

10.15 Statistical Analysis

Data was collated in an Excel[®] database (Microsoft, Redmond, Washington, USA). Statistical analysis was carried out with SPSS[®] software statistical package (SPSS, Chicago, Illinois, USA). $P < 0.05$ was regarded as statistically significant.

Inter-rater reliability analysis between the MSc Students could not be done as there was a large amount of variance within the student group and there were multiple assessors.

10.16 Results

Generic & Specific Ratings of Operative Skill

Table 10.5 summaries the mean scores of the MSc group collectively as well as those for the experienced surgeons for the relevant categories in assessing the technical skills of the operation. Full individual MSc assessors' scores are given in the (Table 10.30). The range for all categories varied largely for the MSc students.

Table 10.5 – MSc Students & Experienced Surgeons Mean Scores

Observer Group	Likert Generic Score & Range	Likert Specific Score & Range	Checklist Generic Score & Range	Checklist Specific Minor Score & Range	Checklist Specific Major Score & Range	Checklist Specific Significant Major Score & Range
MSc	24.1 18-28	47.2 34-52	11.5 10-13	20.5 16-24	28.6 26-32	24.0 24-24
Experienced Surgeons	25.0 24-26	47.0 46-48	11.0 11-11	21.0 20-22	28.0 28-28	24.0 24-24

Table 10.30 – MSc Students & Experienced Surgeons Technical Skills Assessment Scores

Observer Type 2 = MSc 4 = Experienced Surgeons	Individual Assessor's Code	Likert Generic Score	Likert Specific Score	Checklist Generic Score	Checklist Specific Minor Score	Checklist Specific Major Score	Checklist Specific Significant Major Score
2.00	50.00	23.00	48.00	11.00	22.00	30.00	24.00
2.00	51.00	26.00	52.00	12.00	22.00	30.00	24.00
2.00	52.00	20.00	34.00	12.00	22.00	28.00	24.00
2.00	53.00	28.00	52.00	11.00	16.00	28.00	24.00
2.00	54.00	26.00	42.00	12.00	22.00	30.00	24.00
2.00	55.00	18.00	48.00	11.00	22.00	26.00	24.00
2.00	56.00	24.00	50.00	11.00	20.00	28.00	24.00
2.00	57.00	24.00	42.00	10.00	22.00	28.00	24.00
2.00	58.00	24.00	48.00	13.00	18.00	32.00	24.00
2.00	59.00	26.00	48.00	11.00	22.00	28.00	24.00
2.00	60.00	22.00	50.00	13.00	20.00	28.00	24.00
2.00	61.00	24.00	48.00	11.00	18.00	28.00	24.00
2.00	62.00	24.00	52.00	12.00	18.00	28.00	24.00
2.00	63.00	26.00	46.00	11.00	20.00	28.00	24.00
2.00	64.00	25.00	48.00	10.00	22.00	28.00	24.00
2.00	65.00	24.00	48.00	11.00	24.00	30.00	24.00
2.00	66.00	24.00	48.00	13.00	20.00	28.00	24.00
2.00	67.00	26.00	48.00	12.00	20.00	30.00	24.00
4.00	151.00	24.00	48.00	11.00	22.00	28.00	24.00
4.00	152.00	26.00	46.00	11.00	20.00	28.00	24.00

Percentage of MSc Students Agreement to Experienced Surgeons

This was done to determine how accurate individually the MSc Students were to the experienced surgeons in the scoring of technical skills. The mean scores of the experienced surgeons were compared with 95% accuracy (standard deviation +/- 5%) with those of individual MSc students (Table10.6). The accuracy rates for the MSc students were higher than the BSc students with training (Table 10.2). However the accuracy for the MSc students were still not to an acceptable level. They would probably benefit from training to increase their accuracy compared to the experienced surgeons.

Table 10.6 – Percentage Agreement of MSc Students to Experienced Surgeons

Likert Generic Score	Likert Specific Score	Checklist Generic Score	Checklist Specific Minor Score	Checklist Specific Major Score	Checklist Specific Significant Major Score
72%	56%	83%	72%	89%	100%

In summary the MSc students had less of a variance in scores compared to the BSc students. As a group, the MSc students had a better accuracy and closer to the experienced surgeons than the BSc students without training.

10.17 Conclusions

Assessing technical skills in a structured manner in this new era of surgery may be beneficial for training, self-appraisal and re-validation. The type of reliable and accurate assessor must be defined before such tools are implemented. Therefore in these studies in this present chapter we have aimed to evaluate different types of assessors with varying surgical experience in an attempt to answer this question.

In the descriptive statistics, variance in the mean scores in the likert scoring (generic and specific) system compared to the experienced surgeons decreases from the BSc to MSc students. This is because the scale is subjective and requires some surgical experience to use this scale effectively. Therefore the BSc students with the least surgical experience mean scores vary much more than the MSc students who have more surgical experience. The variance in the mean checklists scoring (generic, minor, major, significant major) is much less than that in the likert scoring in each student group compared to the

experienced surgeons. This is probably due to the fact the likert scoring is not a finite scoring system, unlike the checklist scoring, and needs a certain level of subjectivity.

In the percentage agreement statistics the BSc students improved in their accuracy in assessing the generic and specific technical skills after training. The MSc students had a better baseline accuracy rate compared to untrained BSc students to assess each technical skills category. The MSc group could probably benefit from targeted training in technical skills assessment. The MSc group had some surgical experience compared to the BSc group, and this may have been a factor in the percentage agreement rates.

These present studies shows that too many observers/assessors with little surgical experience to assess the technical skills of an operation can not provide good inter-rater reliability, as there is too much variance in the assessment scores, although this improves with structured training.

Conclusions & Applications

Chapter 11

11.10 Overview

Assessing surgical technical skills in a structured manner is a topical issue in light of re-validation, shorter higher surgical training periods and high profile medico-legal cases. The surgical working and training environment is entering a new era in the U.K. Reduction in working hours for surgeons in training, will have a dramatic effect with much less exposure to high volumes of operative work than previous, and the total time for training will be drastically reduced. It is therefore increasingly important that opportunities to acquire technical skills be used as effectively as possible.

The NHS is now a Consultant lead organization which increases the safety of patient care by increasing the level of supervision of surgical trainees in the operating theatre. This prevents the traditional style of acquiring surgical technical skills by trainees practicing on patients. This change however prevents trainees to acquire and perform technical tasks independently, which is an important step in gaining confidence in acquiring the specific technical skills of an operation. There is little research on the effect of implementing a curriculum of acquiring technical skills in surgery with feedback.

In this final chapter we will overview the findings from the studies in the thesis, discuss the limitations of the findings, and applications of the findings to future surgical research, clinical and educational implementations.

11.11 Findings from the Studies

In this thesis we have developed new validated and reliable tools which can assesses both generic and operation specific technical skills in open and laparoscopic surgery. We have also evaluated the benefit of training to assess technical skills, as well as determining what type of surgical experience is required to assess technical skills in surgery.

Hierarchical task analysis was constructed on three key operations, elective primary inguinal hernia repair, sapheno-femoral junction ligation and laparoscopic cholecystectomy. Likert and checklist scales were then constructed individually for generic and operation specific technical skills for all of the operations. Two surgeons with more than 12 years postgraduate surgical experience assessed each operation blindly and independently.

195 operations were analyzed in total, 100 live laparoscopic cholecystectomies, 48 live primary inguinal hernias and 47 (live and simulated) sapheno-femoral junction ligations. These were performed by a total of 17 Consultant Surgeons and 34 Registrar Trainee Surgeons.

Reliability between the observers and construct validity between the two surgeons groups were demonstrated for generic and specific skills in laparoscopic cholecystectomies and inguinal hernia repair using the likert scoring.

Reliability between the observers was good for likert scoring of generic and specific technical skills in live sapheno-femoral junction ligation, however construct validity was not demonstrated for generic and specific technical skills using likert scoring between the Consultant Surgeons and Senior Registrar Surgeons groups. Combining the live and simulated sapheno-femoral junction ligations demonstrated construct validity for generic technical skills only between the Consultant Surgeons and all the Registrar Trainee Surgeons (Junior and Senior).

Regarding assessing technical skills, there is a learning effect by the BSc students who improve with training in assessing the technical skills in laparoscopic cholecystectomies. The MSc students had better baseline knowledge of assessing technical skills compared to the untrained BSc students. Also only a surgeon with adequate surgical and technical experience can reliably and independently assess the generic and specific technical skills in laparoscopic cholecystectomies.

11.12 Limitations of Studies

I will now outline limitations of the studies in this thesis under four themes: Operations, Subjects, Level of Supervision, Level of Surgical Experience and Reality of Operations.

Operations

3 operations were analysed, and they were not complex major operations. There were two open operations chosen and one laparoscopic operation. The operations chosen are index operations and are a good starting point to develop an assessment tool in surgery. This

can then be applied to more complex major operations. The number of operations within each operation group varied, laparoscopic cholecystectomy (100), inguinal hernia (48) and sapheno-femoral junction ligation (47). However this thesis contains a large amount of live operations being assessed compared to any previous study. There were more experts in laparoscopic cholecystectomy group than the open operation group in developing the hierarchical task analysis. This does not affect the hierarchical task analysis for each operation as each is a consensus of experts and is descriptive rather than prescriptive task analysis.

Subjects

There were unequal numbers of Consultants compared to Trainees performing the operations. There were more Consultants in the following groups: Open Live Inguinal Hernia Repair, Open Live Sapheno-Femoral Junction Ligation and Laparoscopic Cholecystectomy. Not all of the Consultant Surgeons performed the same number of operations each within each group.

There were more trainees in the Open Live & Synthetic Sapheno-Femoral Junction Ligation groups compared to Consultants. As a result there were insufficient numbers in this Trainee group for any valuable validity statistical analysis.

Although there were unequal numbers in each operation group, the difference in the technical skills between Trainees and Consultants in the Open Live Inguinal Hernia

Repair and Laparoscopic Cholecystectomy was significant despite these numbers. A larger number of subjects would only further validate the results.

Level of Supervision

In the following groups the trainees were not supervised: Open Live Inguinal Hernia Repair, Open Live Sapheno-Femoral Junction Ligation, Open Live Sapheno-Femoral Junction Ligation. However in the Laparoscopic Cholecystectomy group all of the Trainees were supervised, so there was an influence by the Consultant Surgeons on the Trainees. Therefore the Trainees were prevented from making any serious errors. This later fact does not affect the validity of the assessment tool, as there were very few major or significant major errors committed by both group of surgeons. The other categories being assessed in the technical skills were able to discriminate between levels of experience.

Level of Surgical Experience

In the Open Live Sapheno-Femoral Junction Ligation group the Senior Trainees were almost of Consultant Surgeon status; hence this is why there was no construct validity between the two groups. The two groups were in essence one group regarding experience and ability in technical skills. However if a junior group of Trainees were recruited a difference would most likely be seen.

In the assessors of technical skills the novice surgeons and BSc students had not performed operations by themselves, and did not have the technical skills know how. The

students also vary in their knowledge of the operation. However despite this fact the BSc were able to learn from the training how to assess technical skills in laparoscopic cholecystectomy, and the MSc students as a *group* were comparable to the two experienced surgeons in assessing the technical skills.

Reality of Operations

There were only 20 simulated operations compared to 175 live operations. The simulated operations were not as realistic as performing real operations. However the same generic and specific technical skills are required to perform the operations and therefore does not affect the assessment of technical skills.

11.13 Research

As a result of the research findings from this thesis the following research projects could be conducted. This is a result of the methodology as well as the technical skills assessment tools developed.

Assessing Technical Skills in Other Operations

The template to assess technical skills in the studies in this thesis could be adapted to other operations in different surgical specialties. These operations need different aspects of technical skills to be evaluated. Also this template could be adapted to more complex operations needing advanced specialized technical skills. The following surgical specialties and complex general surgical operations could be evaluated.

A/ Surgical Specialties

- Gynaecology
- Urology
- Orthopaedics
- Cardiothoracic
- Ear, Nose & Throat
- Ophthalmology
- Plastic
- Neurosurgery
- Maxillofacial

B/ Complex General Surgery

- Laparoscopic Colectomy
- Abdominal Aortic Aneurysm Repair
- Oesophagectomy

I now will briefly outline a future research project which can be conducted as a result of this thesis from the above list: Assessing Technical Skills in Complex General Surgery

Assessing the Technical Skills in Complex Surgery - Laparoscopic Colectomy

Malignant and benign bowel disease is one of the most common diseases in the world affecting humans. When patients with significant bowel disease need operative intervention, the normal procedure is to remove the bowel through an open procedure. An open procedure requires a large incision of the abdomen to remove the relevant part of bowel. This results in the patient staying in hospital for a week due to the pain of the wound.

However in Europe and the USA there has been a trend to remove the bowel through small abdominal incisions (keyhole or laparoscopic surgery), which reduces the inpatient stay to 3 days, due to less wound pain. Patients also return to work and normal activities earlier, thus improving their quality of life. In the UK there has been a slow uptake to

learn how to do this keyhole operation. As a consequence there are few surgeons in the UK who perform this procedure regularly in high numbers.

At present there is no formal training or assessment to learn this new surgical technique. Most surgeons learn the technique by watching and going on a surgical skills course, with no structured assessment or feedback of their skills. Subsequently they will learn the operation on real patients. There is a surgical learning curve, which may invariably affect patient safety, with possible poor outcome or adverse events. Recently there is an initiative to implement preceptorships for those who want to learn this new technique.

This research application project would characterize the technical skills and exact steps required to perform this keyhole procedure in a structured way. It will also quantify and assess the significant intra-operative decisions required to satisfactorily complete the operation, as well as how to avoid or deal with difficulties while doing the keyhole operation. Subsequently this may become a template for teaching future surgeons to learn how to perform this advanced surgical procedure satisfactorily in a structured way.

As a result keyhole bowel surgery could be performed more widely in the UK with a structured assessment of the training and technical skills required of the surgeon. This will enable more patients in the UK having this keyhole procedure; they will also therefore have shorter stays in hospital compared to the open procedure. The patients will also be able to return to normal activities and work earlier (if still eligible). This will improve the quality of life of the patients. By reducing the hospital inpatient stay this will

increase the amount of money available to the NHS for other needy concerns, which would have otherwise been spent on these patients staying in hospital longer. Also by returning to normal activities and possibly work the patients will improve his or her psychology, financial status, as well as their local economy.

The main steps of this project would be:

- 1/ Developing a Valid Hierarchical Task Analysis in Laparoscopic Colectomies
- 2/ Develop a Reliable & Valid Technical Skills Assessment Tool in Laparoscopic Colectomies
- 3/ Pilot Studies to Assess Practicality & Feasibility of the Assessment Tool
- 4/ Recruitment of Surgeons Performing Laparoscopic Colectomies from Different Centres of Excellence

11.14 Implementations for Clinical Practice

Reducing Technical Skills Errors in Surgery

From this thesis the execution of technical skills errors has been shown in open and laparoscopic surgery by Trainees as well as Consultant Surgeons. There are various ways to reduce technical skills errors in surgery.

They could be reduced by the following measures:

- 1/ Reduce Complexity
- 2/ Standardize Procedures
- 3/ Checklists and Reminders
- 4/ Standardize Equipment & Instruments
- 5/ Adequate Surgical Training

I will briefly outline two possible future applications to the research findings of this thesis from the above list regarding technical skills errors: Standardizing Operations, Standardizing Equipment & Instruments

Reducing Technical Skills Errors in Surgery

The research findings from this thesis could be used to reduce technical skills errors in open and laparoscopic surgery, and therefore a possible useful clinical application.

Standardize Operations

Comparable surgical operations may be performed differently at different hospitals and even performed differently between two surgeons working within one hospital. Because of this, detailed protocols on how a certain surgical operation or procedure should be precisely performed do not exist. This causes limitations with regard to training of junior surgeons, and it also makes it difficult for the staff in the operating theatre to develop a routine. Clear protocols for every operation could be developed. One way of standardizing a procedure is by using hierarchical task analysis for each operation, in the

form of surgical scripts. This could be combined with pictorial displays, like in a surgical atlas, which would provide for the first time a detailed didactic display of the operations being learnt. All surgeons would follow an agreed pathway of how to perform the operation. The hierarchical task analysis could be a learning tool for trainees, and could prevent or reduce the occurrence of technical skills errors.

Standardize Instruments & Equipment

In laparoscopic surgery the surgeon can use several types of instruments. Different laparoscopic graspers can have different mechanical configurations, leading to different force transmission characteristics. Since the appearance of the instruments may be similar, these differences may not be obvious to the surgeon. In this thesis it has been shown that that surgeons can vary in their surgical techniques using different instruments in laparoscopic cholecystectomies. Some techniques may have more technical error rates compared to others. Standardization of instruments may solve this problem. Another possible aspect of standardizing equipment in open surgery is for all colonic anastomoses to be done by staples. This takes away the human factor aspect of performing a hand sewn anastomosis which can be subject to human error, and therefore leaking anastomoses which can lead to significant morbidity and mortality.

11.15 Education

Implementing a Curriculum of Acquiring Technical Skills

There is little research on the effect of implementing a curriculum of acquiring technical skills in surgery with feedback. The three key groups who would benefit from such a curriculum are:

Basic Surgical Trainees

Higher Surgical Trainees

Surgical Practitioners

In this educational research project application, the aim would be to implement and develop a curriculum of acquiring technical skills in open and laparoscopic surgery by basic and higher surgical trainees in General Surgery and compare it to a control group of basic and higher surgical trainees in General Surgery without a structured curriculum of acquiring technical skills. The main aims of this future application are:

- 1/ Develop a curriculum for basic and higher surgical trainees in General Surgery in open and laparoscopic operations and assess their technical skills.
- 2/ Assess technical skills in basic and higher surgical trainees in General Surgery in open and laparoscopic procedures without a structured curriculum of acquiring technical skills in the era of reduced working hours.
- 3/ Provide a framework with continuous feedback within which curriculum trainees can use to improve their technical skill longitudinally

This future possible educational application could be initially in a form of a randomised study, as outlined in Figures 11.1, 11.2, 11.3 and 11.4. Such an implementation curriculum will demonstrate whether a structured curriculum of acquiring technical skills is beneficial to basic and higher surgical trainees in this evolving surgical training and work environment. This educational surgical programme would enable trainee surgeons to acquire generic and specific technical skills in a shorter period of time, and therefore

complement the shorter period of surgical training which is currently being implemented in the UK.

Self- Appraisal of Surgical Technical Skills

The General Medical Council (GMC) is implementing compulsory re-validation for every doctor, every five years. This is mainly a paper process; however for surgeons it does not evaluate their operative technique or skills. The ability to assess a surgeon's technical skills in a structured manner would be beneficial for this process.

Technical skills self-appraisal had been used previously using video tapes of operations by Trainees/Residents and Consultants/Attending, and the surgeon evaluating his or her technical skills performance using a structured assessment criteria. However to date there has been no analysis of detailed self-appraisal of expert surgeons.

A prospective, objective method of analyzing surgical technical skills errors can be used, in the shape of Hierarchical Task Analysis (HTA) or using the structured assessment of technical skills developed in this thesis. Experts and trainees could use these tools to self-appraise themselves. They could do this after the operation or, if the operation has been recorded, to do this using video footage.

11.16 Summary

This thesis has attempted for the first time to assess comprehensively the generic and specific technical skills of surgeons with varying experience in open and laparoscopic surgery. The integral aspects of this thesis can be applied to various different operations as well as other surgical educational and clinical applications. The research findings from this thesis can be applied to further research projects related to surgical competency and training.

It can be seen that the research findings of this thesis has possible applications to the daily performance of surgery as well as a role in the education of the surgeons of the future. Learning and assessing technical skills in a structured manner in expert and trainee surgeons is beneficial for the surgical community as well as for the rest of the wider community. By developing these possible frameworks in the research, educational and clinical domains in surgery, patients and surgeons would benefit and therefore ultimately improve surgical patient safety.

Figure 11.1 - Basic Surgical Trainee – Curriculum of Assessment & Feedback



Figure 11.2 - Higher Surgical Trainee – Curriculum of Assessment & Feedback

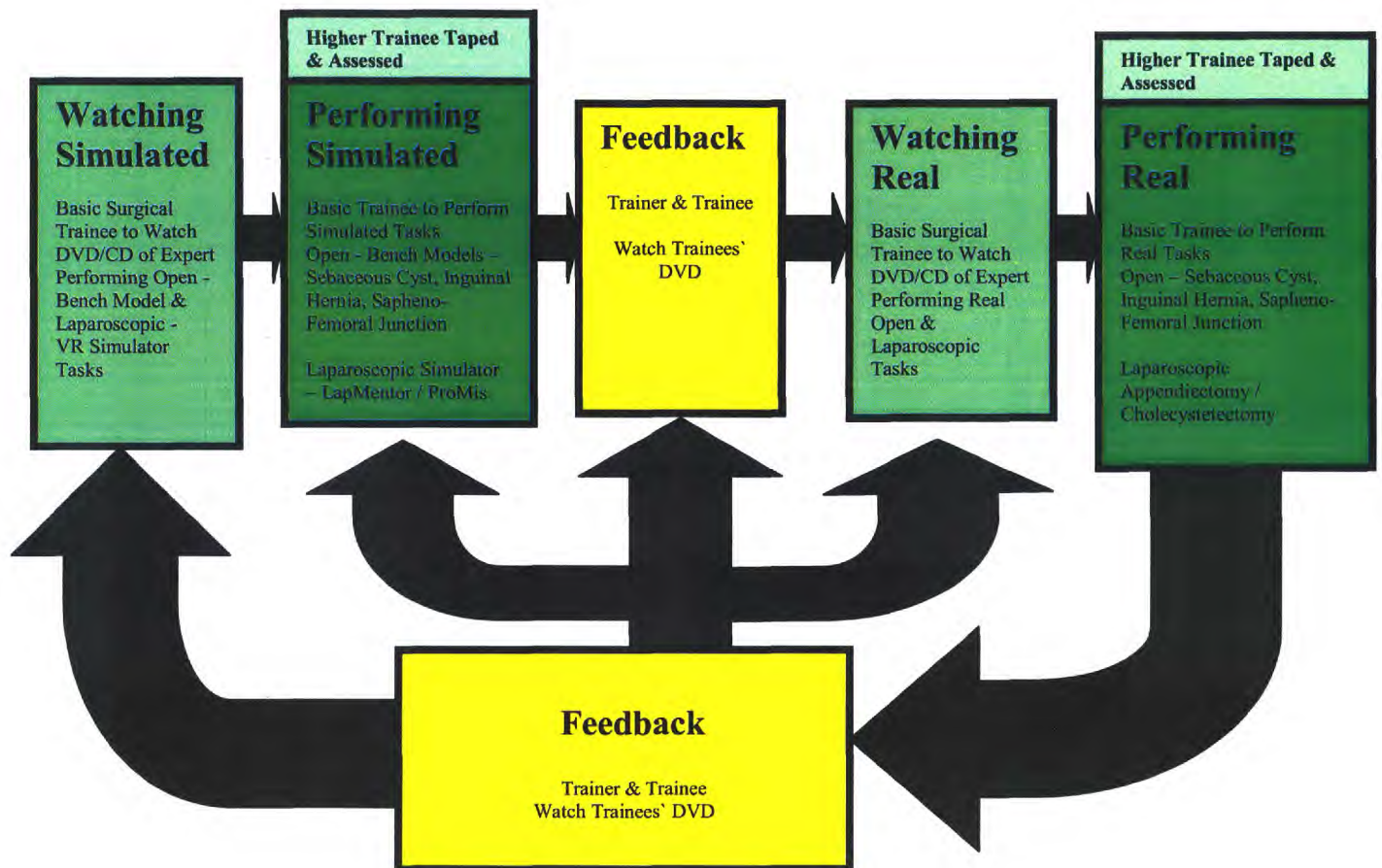


Figure 11.3 - Basic Surgical Trainee – No Curriculum of Assessment & Feedback

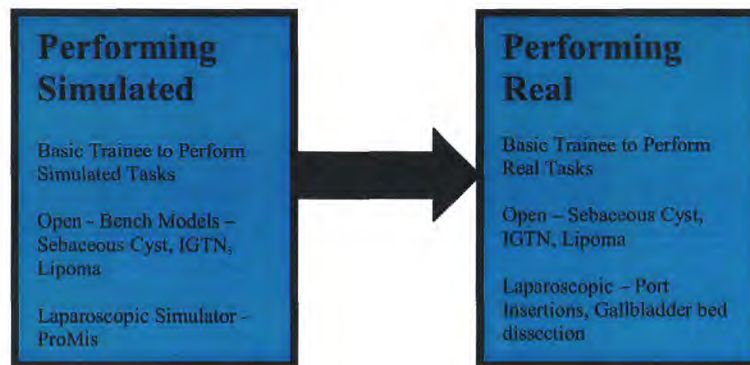
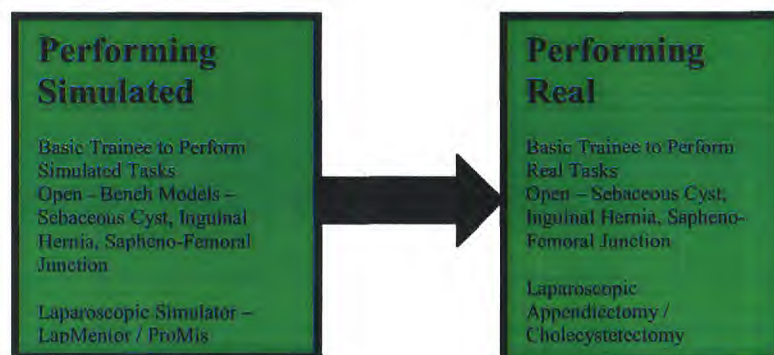


Figure 11.4 - Higher Surgical Trainee - No Curriculum of Assessment & Feedback



References

- Adamsen S, Funch-Jensen PM, Drewes AM, Rosenberg J, Grantcharov TP.
A comparative study of skills in virtual laparoscopy and endoscopy.
Surg Endosc. 2005 Feb;19(2):229-34.
- Adrales GL, Park AE, Chu UB, Witzke DB, Donnelly MB, Hoskins JD, Mastrangelo MJ Jr, Gandsas A.
A valid method of laparoscopic simulation training and competence assessment.
J Surg Res. 2003 Oct;114(2):156-62.
- Ahlberg G, Heikkinen T, Iselius L, Leijonmarck CE, Rutqvist J, Arvidsson D
Does training in a virtual reality simulator improve surgical performance?
Surg Endosc. 2002 Jan;16(1):126-9.
- Albayrak A, Casseres YA, Schot C, Grimbergen CA, Bonjer HJ, Meijer D
Quality assessment of light guide cables and endoscopes in Dutch hospitals.
Min Invas Ther Allied Technol 2003 12: 179
- Alves A, Panis Y, Slim K, Heyd B, Kwiatkowski F, Manton G; Association Francaise de Chirurgie.
French multicentre prospective observational study of laparoscopic versus open colectomy for sigmoid diverticular disease.
Br J Surg. 2005 Dec;92(12):1520-5
- Amid PK, Shulman AG, Lichtenstein IL.
Open "tension-free" repair of inguinal hernias: the Lichtenstein technique.
Eur J Surg. 1996 Jun;162(6):447-53
- Anastakis DJ¹, Wanzel KR, Brown MH, McIlroy JH, Hamstra SJ, Ali J, Hutchison CR, Murnaghan J, Reznick RK, Regehr G.
Evaluating the effectiveness of a 2-year curriculum in a surgical skills center.
Am J Surg. 2003 Apr;185(4):378-85.

Anastakis DJ ², Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, Hutchison C.

Assessment of technical skills transfer from the bench training model to the human model.

Am J Surg. 1999 Feb;177(2):167-70.

Anonymous, Irving M, Berwick DM, Treasure T.

Five times: coincidence or something more serious?

BMJ 1998;316:1736–1740.

Armitage P, Berry G, Matthews J N S

Statistical Methods in Medical Research

Fourth Edition

2002 Blackwell Publishing

Arora H, Uribe J, Ralph W, Zeltsan M, Cuellar H, Gallagher A, Fried MP.

Assessment of construct validity of the endoscopic sinus surgery simulator.

Arch Otolaryngol Head Neck Surg. 2005 Mar;131(3):217-21.

Ault G, Reznick R, MacRae H, Leadbetter W, DaRosa D, Joehl R, Peters J, Regehr G.

Exporting a technical skills evaluation technology to other sites.

Am J Surg. 2001 Sep;182(3):254-6.

Backstein ¹ D, Agnidis Z, Regehr G, Reznick R. The effectiveness of video feedback in the acquisition of orthopedic technical skills.

Am J Surg. 2004 Mar;187(3):427-32.

Backstein ² D, Agnidis Z, Sadhu R, MacRae H.

Effectiveness of repeated video feedback in the acquisition of a surgical technical skill.

Can J Surg. 2005 Jun;48(3):195-200.

Bann S ¹, Kwok KF, Lo CY, Darzi A, Wong J.

Objective assessment of technical skills of surgical trainees in Hong Kong.

Br J Surg. 2003 Oct;90(10):1294-9.

Bann S ², Datta V, Khan M, Darzi A.

The surgical error examination is a novel method for objective technical knowledge assessment.

Am J Surg. 2003 Jun;185(6):507-11.

Bann SD ³, Khan MS, Darzi AW.

Measurement of surgical dexterity using motion analysis of simple bench tasks.

World J Surg. 2003 Apr;27(4):390-4.

Bann S ⁴, Khan MS, Datta VK, Darzi AW.

Technical performance: relation between surgical dexterity and technical knowledge.

World J Surg. 2004 Feb;28(2):142-6; discussion 146-7.

Bann S ⁵, Khan M, Datta V, Darzi A.

Surgical skill is predicted by the ability to detect errors.

Am J Surg. 2005 Apr;189(4):412-5.

Bann S ⁶, Davis IM, Moorthy K, Munz Y, Hernandez J, Khan M, Datta V, Darzi A.

The reliability of multiple objective measures of surgery and the role of human performance.

Am J Surg. 2005 Jun;189(6):747-752.

Barach P., Small S.D.

Reporting and preventing medical mishaps:

lessons from non-medical near miss reporting systems

BMJ 2000; 320: 759-763

Barnes RW ¹.

Surgical handicraft: teaching and learning surgical skills.

Am J Surg. 1987 May;153(5):422-7.

Barnes RW ², Lang NP, Whiteside MF.

Halstedian technique revisited. Innovations in teaching surgical skills.

Ann Surg. 1989 Jul;210(1):118-21.

Beard ¹ JD, Jolly BC, Newble DI, Thomas WE, Donnelly J, Southgate LJ.

Assessing the technical skills of surgical trainees.

Br J Surg. 2005 Jun;92(6):778-82

Beard ² JD, Jolly BC, Southgate LJ, Newble DI, Thomas EG, Rochester J.

Developing assessments of surgical skills for the GMC Performance Procedures.

Ann R Coll Surg Engl. 2005 Jul;87(4):242-7

Begg CB, Riedel ER, Bach PB, Kattan MW, Schrag D, Warren JL, Scardino PT.

Variations in morbidity after radical prostatectomy.

N Eng J Med 2002;346:1138-44.

Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, Welch

HG, Wennberg DE.

Hospital volume and surgical mortality in the United States.

N Eng J Med 2002;346:1128-37.

Bland J.M.

An Introduction to Medical Statistics

Third Edition

2000 Oxford Publishing, Oxford

Bogner SM (Ed)

Human Error in Medicine

1994 Lawrence Erlbaum Associates, Hillsdale, New Jersey, USA

Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers

AG, Newhouse JP, Weiler PC, Hiatt HH

Incidence of adverse events and negligence in hospitalized patients: results of the

Harvard Medical Practice Study I.

N Engl J Med 1991; 324:370-376

Brewster LP, Risucci DA, Joehl RJ, Littooy FN, Temeck BK, Blair PG,

Sachdeva AK.

Management of adverse surgical events: a structured education module for residents.

Am J Surg. 2005 Nov;190(5):687-90

Burt ¹ C.G. Ricketts C. Grant J.R. Wilkins D.C.

The validity of the operative competence assessment for surgical trainees

Br J Surg 2005; 92; s111

Burt ² C.G. Wilkins D.C.

The accuracy of self-assessment by surgical trainees

Br J Surg 2005; 92; s140

Calland JF, Guerlain S, Adams RB et al

A systems approach to surgical safety.

Surg Endosc 2002;1005-14.

Cao ¹ C.G.L., MacKenzie C.L, Payandeh S.

Conference Proceedings: 5th Annual Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. 1996, 583-590

Cao ² CGL, MacKenzie CL, Ibbotson JA, Turner LJ, Blair NP, Nagy AG

Hierarchical Decomposition of Laparoscopic Procedures. Medicine Meets Virtual Reality, 1999 IOS Press

Carter D. The surgeon as a risk factor BMJ 2003; 326: 832-3

Cass OW, Freeman ML, Peine CJ, Zera RT, Onstad GR.

Objective evaluation of endoscopy skills during training.

Ann Intern Med. 1993 Jan 1;118(1):40-4.

Casseres YA, Albayrak A, Schot C, Grimbergen CA, Bonjer HJ, Meijer D

The incidence of insulation failure of endoscopic instruments in Dutch hospitals.

Min Invas Ther Allied Technol 2003 12: 178

Cauraugh JH, Martin M, Martin KK.

Modeling surgical expertise for motor skill acquisition.

Am J Surg. 1999 Apr;177(4):331-6.

Chak A, Cooper GS, Blades EW, Canto M, Sivak MV Jr.

Prospective assessment of colonoscopic intubation skills in trainees.

Gastrointest Endosc. 1996 Jul;44(1):54-7.

Chatenay M, Maguire T, Skakun E, Chang G, Cook D, Warnock GL.

Does volume of clinical experience affect performance of clinical clerks on surgery exit examinations?

Am J Surg. 1996 Oct;172(4):366-72.

Chetter IC, Mylankal KJ, Hughes H, Fitridge R.

Randomized clinical trial comparing multiple stab incision phlebectomy and transilluminated powered phlebectomy for varicose veins.

Br J Surg. 2006 Feb;93(2):169-74.

Cohen MM, Young W, Theriault ME, Hernandez R.

Has laparoscopic cholecystectomy changed patterns of practice and patient outcome in Ontario?

CMAJ. 1996 Feb 15;154(4):491-500.

Coleman RL, Muller CY.

Effects of a laboratory-based skills curriculum on laparoscopic proficiency: a randomized trial.

Am J Obstet Gynecol. 2002 Apr;186(4):836-42.

Connor S, Garden OJ.

Bile duct injury in the era of laparoscopic cholecystectomy.

Br J Surg. 2006 Feb;93(2):158-68.

Copeland¹ G P, Jones D, Walters M.

POSSUM: a scoring system for surgical audit.

Br J Surg 1991; 78: 355-360.

Copeland² G P.

The POSSUM system of surgical audit.

Arch Surg 2002; 137: 15-19.

Crawford SW, Colt HG.

Virtual reality and written assessments are of potential value to determine knowledge and skill in flexible bronchoscopy.

Respiration. 2004 May-Jun;71(3):269-75.

Cuschieri ¹ A.

Reflections on surgical training.

Surg Endosc. 1993 Mar-Apr;7(2):73-4.

Cuschieri ² A.

Reducing errors in the operating room: surgical proficiency and quality assurance of execution.

Surg Endosc. 2005 Aug;19(8):1022-7

Dankelman J, Grimbergen CA.

Systems approach to reduce errors in surgery.

Surg Endosc. 2005 Aug;19(8):1017-21

Darzi A, Datta V, Mackay S.

The challenge of objective assessment of surgical skill

Am J Surg 2001;181:484-6.

Datta V ¹, Mackay S, Mandalia M, Darzi A.

The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model.

J Am Coll Surg. 2001 Nov;193(5):479-85.

Datta V ², Chang A, Mackay S, Darzi A.

The relationship between motion analysis and surgical technical assessments.

Am J Surg. 2002 Jul;184(1):70-3.

Datta V ³, Mandalia M, Mackay S, Darzi A.

The PreOp flexible sigmoidoscopy trainer. Validation and early evaluation of a virtual reality based system.

Surg Endosc. 2002 Oct;16(10):1459-63

Datta V ⁴, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A.
Relationship between skill and outcome in the laboratory-based model.
Surgery. 2002 Mar;131(3):318-23.

Datta V ⁵, Bann S, Beard J, Mandalia M, Darzi A.
Comparison of bench test evaluations of surgical skill with live operating
performance assessments.
J Am Coll Surg. 2004 Oct;199(4):603-6.

Den Boer ¹ KT, De Wit LT, Dankelman J, Gouma DJ.
Peroperative time-motion analysis of diagnostic laparoscopy with laparoscopic
ultrasonography.
Br J Surg. 1999 Jul;86(7):951-5.

Den Boer ² KT, Dankelman J, Gouma DJ, Stassen HG.
Peroperative analysis of surgical procedure
Surg Endosc. 2002;16: 492-499

Derossis AM, Antoniuk M, Fried GM.
Evaluation of laparoscopic skills: a 2-year follow-up during residency training.
Can J Surg. 1999 Aug;42(4):293-6.

De Leval MR, Francois K, Bull C, Brawn W, Spiegelhalter D. Analysis of a cluster of
surgical failures. Application to a series of neonatal arterial switch operations. J
Thoracic Cardiovasc Surg. 1994;107:914-924.

Donabedian A.
Promoting quality through evaluating the process of patient care.
Med Care 1968;6:181-201

Dosis ¹ A, Bello F, Moorthy K, Munz Y, Gillies D, Darzi A.
Real-time synchronization of kinematic and video data for the comprehensive
assessment of surgical skills.
Stud Health Technol Inform. 2004;98:82-8.

- Dosis ² A, Aggarwal R, Bello F, Moorthy K, Munz Y, Gillies D, Darzi A.
Synchronized video and motion analysis for the assessment of procedures in the operating theater.
Arch Surg. 2005 Mar;140(3):293-9
- Dubrowski A, Sidhu R, Park J, Carnahan H.
Quantification of motion characteristics and forces applied to tissues during suturing.
Am J Surg. 2005 Jul;190(1):131-6
- Eastridge BJ, Hamilton EC, O'Keefe GE, Rege RV, Valentine RJ, Jones DJ, Tesfay S, Thal ER.
Effect of sleep deprivation on the performance of simulated laparoscopic surgical skill.
Am J Surg. 2003 Aug;186(2):169-74.
- Emam TA, Hanna GB, Kimber C, Cuschieri A. Differences between experts and trainees in the motion pattern of the dominant upper limb during intracorporeal endoscopic knotting.
Dig Surg 2000;17: 120-3.
- Embrey D, 2000. Task Analysis Techniques, Human Reliability Associates Ltd
- Eubanks TR, Clements RH, Pohl D, Williams N, Schaad DC, Horgan S, Pellegrini C.
An objective scoring system for laparoscopic cholecystectomy.
J Am Coll Surg. 1999 Dec;189(6):566-74.
- Eriksen JR, Grantcharov T.
Objective assessment of laparoscopic skills using a virtual reality stimulator.
Surg Endosc. 2005 Sep;19(9):1216-9
- Eversbusch A, Grantcharov TP.
Learning curves and impact of psychomotor training on performance in simulated colonoscopy: a randomized trial using a virtual reality endoscopy trainer.
Surg Endosc. 2004 Oct;18(10):1514-8.

Faulkner H, Regehr G, Martin J, Reznick R.

Validation of an objective structured assessment of technical skill for surgical residents.

Acad Med. 1996 Dec;71(12):1363-5.

Federal Aviation Administration. The aviation safety system. www.faa.gov

Federal Aviation Administration (1995).

Human performance considerations in the use and design of aircraft checklists.

www.faa.gov/avr/afs/checklist.doc

Feldman LS, Hagarty SE, Ghitulescu G, Stanbridge D, Fried GM.

Relationship between objective assessment of technical skills and subjective in-training evaluations in surgical residents.

J Am Coll Surg. 2004 Jan;198(1):105-10.

Fletcher DR, Hobbs MS, Tan P, Valinsky LJ, Hockey RL, Pikora TJ, Knuiman MW, Sheiner HJ, Edis A.

Complications of cholecystectomy: risks of the laparoscopic approach and protective effects of operative cholangiography: a population-based study.

Ann Surg. 1999 Apr;229(4):449-57.

Flum DR, Dellinger EP, Cheadle A, Chan L, Koepsell T.

Intraoperative cholangiography and risk of common bile duct injury during cholecystectomy.

JAMA. 2003 Apr 2;289(13):1639-44.

Francis ¹ NK, Hanna GB, Cresswell AB, Carter FJ, Cuschieri A.

The performance of master surgeons on standard aptitude testing.

Am J Surg. 2001 Jul;182(1):30-3.

Francis ² NK, Hanna GB, Cuschieri A.

Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks.

Arch Surg. 2001 Jan;136(1):40-3.

Francis³ NK, Hanna GB, Cuschieri A.

The performance of master surgeons on the advanced Dundee endoscopic psychomotor tester: contrast validity study.

Arch Surg 2002;137: 841-4.

Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, Andrew CG.

Proving the value of simulation in laparoscopic surgery.

Ann Surg. 2004 Sep;240(3):518-25

Friedlich M, MacRae H, Oandasan I, Tannenbaum D, Batty H, Reznick R, Regehr G. Structured assessment of minor surgical skills (SAMSS) for family medicine residents.

Acad Med. 2001 Dec;76(12):1241-6.

Galasko C, Mackay C.

Unsupervised surgical training. Logbooks are essential for assessing progress.

BMJ. 1997 Nov 15;315(7118):1306-7.

Gallagher AG¹, McClure N, McGuigan J, Crothers I, Browning J.

Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR).

Endoscopy. 1999 May;31(4):310-3.

Gallagher AG², Ritter EM, Satava RM.

principles of validation, and reliability: rigorous science for the assessment of surgical education and training.

Surg Endosc. 2003 Oct;17(10):1525-9.

Gallagher AG³, Smith CD, Bowers SP, Seymour NE, Pearson A, McNatt S, Hananel D, Satava RM.

Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures.

J Am Coll Surg. 2003 Sep;197(3):479-88.

Gallagher AG ⁴, Cowie R, Crothers I, Jordan-Black JA, Satava RM.

PicSOr: an objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopic performance.

Surg Endosc. 2003 Sep;17(9):1468-71.

Gallagher AG ⁵, Lederman AB, McGlade K, Satava RM, Smith CD.

Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance.

Surg Endosc. 2004 Apr;18(4):660-5.

Gettman MT, Kondraske GV, Traxer O, Ogan K, Napper C, Jones DB, Pearle MS, Cadeddu JA.

Assessment of basic human performance resources predicts operative performance of laparoscopic surgery.

J Am Coll Surg. 2003 Sep;197(3):489-96.

Goff BA ¹, Lentz GM, Lee DM, Mandel LS.

Formal teaching of surgical skills in an obstetric-gynecologic residency.

Obstet Gynecol. 1999 May;93(5 Pt 1):785-90.

Goff BA ², Lentz GM, Lee D, Houmard B, Mandel LS.

Development of an objective structured assessment of technical skills for obstetric and gynecology residents.

Obstet Gynecol. 2000 Jul;96(1):146-50.

Goff BA ³, Lentz GM, Lee D, Fenner D, Morris J, Mandel LS.

Development of a bench station objective structured assessment of technical skills.

Obstet Gynecol. 2001 Sep;98(3):412-6.

Goff BA ⁴, Nielsen PE, Lentz GM, Chow GE, Chalmers RW, Fenner D, Mandel LS.

Surgical skills assessment: a blinded examination of obstetrics and gynecology residents.

Am J Obstet Gynecol. 2002 Apr;186(4):613-7.

Goff B ⁵, Mandel L, Lentz G, Vanblaricom A, Oelschlager AM, Lee D, Galakatos A, Davies M, Nielsen P.

Assessment of resident surgical skills: is testing feasible?

Am J Obstet Gynecol. 2005 Apr;192(4):1331-8

Gordon NLM, Dawson AA, Bennett B, Innes G, Eremin O, Jones PF.

Outcome in colorectal adenocarcinoma: two seven-year studies for a population.

BMJ 1993; 307(6906): 707-10

Gosman GG, Simhan HN, Guido RS, Lee TT, Mansuria SM, Sanfilippo JS.

Focused assessment of surgical performance: difficulty with faculty compliance.

Am J Obstet Gynecol. 2005 Nov;193(5):1811-6.

Grober ED ¹, Hamstra SJ, Wanzel KR, Reznick RK, Matsumoto ED,

Sidhu RS, Jarvi KA.

Validation of novel and objective measures of microsurgical skill: Hand-motion analysis and stereoscopic visual acuity.

Microsurgery. 2003;23(4):317-22.

Grober ED ², Hamstra SJ, Wanzel KR, Reznick RK, Matsumoto ED,

Sidhu RS, Jarvi KA.

Laboratory based training in urological microsurgery with bench model simulators: a randomized controlled trial evaluating the durability of technical skill.

J Urol. 2004 Jul;172(1):378-81.

Grober ED ³, Hamstra SJ, Wanzel KR, Reznick RK, Matsumoto ED, Sidhu RS,

Jarvi KA.

The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures.

Ann Surg. 2004 Aug;240(2):374-81.

Grantcharov TP ¹, Bardram L, Funch-Jensen P, Rosenberg J.

Laparoscopic performance after one night on call in a surgical department: prospective study.

BMJ. 2001 Nov 24;323(7323):1222-3.

Grantcharov TP ², Rosenberg J, Pahle E, Funch-Jensen P.

Virtual reality computer simulation.

Surg Endosc. 2001 Mar;15(3):242-4.

Grantcharov TP ³, Bardram L, Funch-Jensen P, Rosenberg J.

Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills.

Am J Surg. 2003 Feb;185(2):146-9.

Grantcharov TP ⁴, Kristiansen VB, Bendix J, Bardram L, Rosenberg J,

Funch-Jensen P.

Randomized clinical trial of virtual reality simulation for laparoscopic skills training.

Br J Surg. 2004 Feb;91(2):146-50.

Haeger K.

The Illustrated History of Surgery

Revised & Updated by Sir Roy Calne

2000 Harold Starke Publishers Ltd, London

Haga Y, Wada Y, Takeuchi H, Kimura O, Furuya T, Sameshima H, Ishikawa M.

Estimation of physiologic ability and surgical stress (E-PASS) for a surgical audit in elective digestive surgery.

Surgery 2004 Jun;135(6):586-94.

Hamilton EC, Scott DJ, Kapoor A, Nwariaku F, Bergen PC, Rege RV,

Tesfay ST, Jones DB.

Improving operative performance using a laparoscopic hernia simulator.

Am J Surg. 2001 Dec;182(6):725-8.

Hammond L, Ketchum J, Schwartz BF.

Accreditation council on graduate medical education technical skills competency compliance: urologic surgical skills.

J Am Coll Surg. 2005 Sep;201(3):454-7.

Hance ¹ J, Aggarwal R, Stanbridge R, Blauth C, Munz Y, Darzi A, Pepper J.
Objective assessment of technical skills in cardiac surgery.
Eur J Cardiothorac Surg. 2005 Jul;28(1):157-62

Hance ² J, Aggarwal R, Moorthy K, Munz Y, Undre S, Darzi A.
Assessment of psychomotor skills acquisition during laparoscopic cholecystectomy
courses.
Am J Surg. 2005 Sep;190(3):507-11

Hanna ¹ GB, Drew T, Cuschieri A.
Technology for Psychomotor Skills Testing in Endoscopic Surgery.
Semin Laparosc Surg. 1997 Jun;4(2):120-124.

Hanna ² GB, Frank TG, Cuschieri A.
Objective assessment of endoscopic knot quality.
Am J Surg. 1997 Oct;174(4):410-3.

Healey MA, Shackford SR, Osler TM, Rogers FB, Burns E.
Complications in surgical patients.
Arch Surg 2002; 137(5): 611-8

Helmreich RL, Schaefer H.
Team performance in the operating room. In Bogner MS,
1994 Ed. Human error in medicine, Hillsdale, New Jersey: Lawrence Erlbaum

Hernandez JD, Bann SD, Munz Y, Moorthy K, Datta V, Martin S, Dosis A, Bello F,
Darzi A, Rockall T.
Qualitative and quantitative analysis of the learning curve of a simulated surgical task
on the da Vinci system.
Surg Endosc. 2004 Mar;18(3):372-8.

Hollnagel, E. Context, cognition, and control. In Y. Waern (Ed.). Co-operation in
process management - Cognition and information technology.
1998 London: Taylor & Francis

Humphreys P. (Ed) Human Reliability Assessors Guide, Safety 7 Reliability
Directorate, Warrington, Publication No. RTS 88/95Q 1995

Jain PK, Hayden JD, Sedman PC, Royston CM, O'Boyle CJ. A prospective study of ambulatory laparoscopic cholecystectomy: training economic, and patient benefits. Surg Endosc. 2005 Aug;19(8):1082-5.

Johnson P

Human-computer interaction: psychology, task analysis and software engineering
1992 1st edition, McGraw-Hill

Joice ¹ P, Hanna GB, Cuschieri A.

Ergonomic evaluation of laparoscopic bowel suturing.
Am J Surg. 1998 Oct;176(4):373-8.

Joice ² P, Hanna GB, Cuschieri A.

Errors enacted during endoscopic surgery--a human reliability analysis.
Appl Ergon. 1998 Dec;29(6):409-14.

Jones H J, de Cossart L.

Risk scoring in surgical patients.
Br J Surg 1999; 86: 149-157.

Kable AK, Gibberd RW, Spigelman AD.

Adverse events in surgical patients in Australia.
Int J Qual Health Care. 2002 Aug;14(4):269-76.

Kariv Y, Delaney CP, Casillas S, Hammel J, Nocero J, Bast J, Brady K, Fazio VW, Senagore AJ.

Long-term outcome after laparoscopic and open surgery for rectal prolapse: a case-control study.
Surg Endosc. 2006 Jan;20(1):35-42.

Kark AE, Kurzer MN, Belsham PA.

Three thousand one hundred seventy-five primary inguinal hernia repairs: advantages of ambulatory open mesh repair using local anesthesia.
J Am Coll Surg. 1998 Apr;186(4):447-55

Katz R, Hoznek A, Salomon L, Antiphon P, de la Taille A, Abbou CC.

Skill assessment of urological laparoscopic surgeons: can criterion levels of surgical performance be determined using the pelvic box trainer?

Eur Urol. 2005 Apr;47(4):482-7.

Khan MS, Bann SD, Darzi A, Butler PE.

Use of suturing as a measure of technical competence.

Ann Plast Surg. 2003 Mar;50(3):304-8

Kirk RM.

Teaching the craft of operative surgery.

Ann R Coll Surg Engl. 1996 Jan;78(1 Suppl):25-8.

Kirwan ¹ B.& L.K. Ainsworth, A Guide to Task Analysis,
Taylor & Francis, London, 1992.

Kirwan ² B., A Guide to Practical Human Reliability Assessment
Taylor & Francis, London, 1994.

Knaus W A, Draper E A.

APACHE II: A severity of disease classification system.

Crit Care Med 1985; 13: 818-829.

Kohls-Gatzoulis JA, Regehr G, Hutchison C.

Teaching cognitive skills improves learning in surgical skills courses: a blinded, prospective, randomized study.

Can J Surg. 2004 Aug;47(4):277-83.

Korndorffer JR Jr, Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ.

Simulator training for laparoscopic suturing using performance goals translates to the operating room.

J Am Coll Surg. 2005 Jul;201(1):23-9

Larson JL, Williams RG, Ketchum J, Boehler ML, Dunnington GL.
Feasibility, reliability and validity of an operative performance rating system for
evaluating surgery residents.
Surgery. 2005 Oct;138(4):640-7

Law WL, Lam CM, Lee YM.
Evaluation of outcome of laparoscopic colorectal resection with POSSUM,
Portsmouth POSSUM and colorectal POSSUM.
Br J Surg. 2005 Nov 15;

Leape ¹ LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, Hebert L,
Newhouse JP, Weiler PC, Hiatt H
The nature of adverse events in hospitalized patients: results of the Harvard Medical
Practice Study II.
N Engl J Med 1991; 324:377-384

Leape ² LL
Institute of medicine: medical error figures are not exaggerated.
J Am Med Assoc 2000; 284:95-97

Liem MS, van der Graaf Y, van Steensel CJ, Boelhouwer RU, Clevers GJ, Meijer
WS, Stassen LP, Vente JP, Weidema WF, Schrijvers AJ, van Vroonhoven TJ.
Comparison of conventional anterior surgery and laparoscopic surgery for inguinal-
hernia repair.
N Engl J Med. 1997 May 29;336(22):1541-7.

Lossing AG, Hatswell EM, Gilas T, Reznick RK, Smith LC.
A technical-skills course for 1st-year residents in general surgery: a descriptive study.
Can J Surg. 1992 Oct;35(5):536-40.

Lunevicius R, Morkevicius M.
Comparison of laparoscopic versus open repair for perforated duodenal ulcers.
Surg Endosc. 2005 Dec;19(12):1565-71.

Lyons M., Adams S., Woloshynowych M., Vincent C.

Human Reliability Analysis in Healthcare: a Review of Techniques

The International Journal of Risk and Safety in Medicine 2004; 16, 223-237.

Mackay S ¹, Datta V, Mandalia M, Bassett P, Darzi A.

Electromagnetic motion analysis in the assessment of surgical skill: Relationship between time and movement.

ANZ J Surg. 2002 Sep;72(9):632-4.

Mackay S ², Morgan P, Datta V, Chang A, Darzi A.

Practice distribution in procedural skills training: a randomized controlled trial.

Surg Endosc. 2002 Jun;16(6):957-61.

Mackay S ³, Datta V, Chang A, Shah J, Kneebone R, Darzi A.

Multiple Objective Measures of Skill (MOMS): a new approach to the assessment of technical ability in surgical trainees.

Ann Surg. 2003 Aug;238(2):291-300.

Macmillan AI, Cuschieri A.

Assessment of innate ability and skills for endoscopic manipulations by the Advanced Dundee Endoscopic Psychomotor Tester: predictive and concurrent validity.

Am J Surg. 1999 Mar;177(3):274-7.

Mahmood T, Darzi A.

A study to validate the colonoscopy simulator.

Surg Endosc. 2003 Oct;17(10):1583-9.

McCloy R, Stone R.

Science, medicine, and the future. Virtual reality in surgery.

BMJ 2001;323: 912-5

McClusky DA 3rd, Ritter EM, Lederman AB, Gallagher AG, Smith CD.

Correlation between perceptual, visuo-spatial, and psychomotor aptitude to duration of training required to reach performance goals on the MIST-VR surgical simulator.

Am Surg. 2005 Jan;71(1):13-20

McMahon AJ, Russell IT, Baxter JN, Ross S, Anderson JR, Morran CG, Sunderland G, Galloway D, Ramsay G, O'Dwyer PJ.

Laparoscopic versus minilaparotomy cholecystectomy: a randomised trial.
Lancet. 1994 Jan 15;343(8890):135-8

MacRae H, Regehr G, Leadbetter W, Reznick RK.

A comprehensive examination for senior surgical residents.
Am J Surg. 2000 Mar;179(3):190-3.

MacDonald J, Williams RG, Rogers DA.

Self-assessment in simulation-based surgical skills training.
Am J Surg. 2003 Apr;185(4):319-22.

McNatt SS, Smith CD.

A computer-based laparoscopic skills assessment device differentiates experienced from novice laparoscopic surgeons.
Surg Endosc. 2001 Oct;15(10):1085-9.

Martin ¹ J., Regehr G., Reznick R. et al

An objective, structured assessment of technical skill for surgical residents
Gastroenterology 1995 108 (Supl): A1231

Martin ² JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M.

Objective structured assessment of technical skill (OSATS) for surgical residents.
Br J Surg. 1997 Feb;84(2):273-8.

Mehta NY, Haluck RS, Frecker MI, Snyder AJ. Sequence and task analysis of instrument use in common laparoscopic procedures.

Surg Endosc. 2002;16: 280-285

Michaels JA, Brazier JE, Campbell WB, Macintyre JB, Palfreyman SJ, Ratcliffe J.

Randomized clinical trial comparing surgery with conservative treatment for uncomplicated varicose veins.

Br J Surg. 2006 Feb;93(2):175-81

Moorthy K ¹, Smith S, Brown T, Bann S, Darzi A.

Evaluation of virtual reality bronchoscopy as a learning and assessment tool.

Respiration. 2003 Mar-Apr;70(2):195-9.

Moorthy K. ², Muntz Y., Sarker S.K., Darzi A.

Objective Assessment of Technical Skills in Surgery

BMJ 2003; 327:1032-7

Moorthy ³ K, Munz Y, Adams S, Pandey V, Darzi A.

A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre.

Ann Surg. 2005 Nov;242(5):631-9

Moossa AR, Mayer AD, Stabile B.

Iatrogenic injury to the bile duct. Who, how, where?

Arch Surg. 1990 Aug;125(8):1028-30

Moray N 1994 Error reduction as a systems problem.

In: Bogner MS (Ed) Human error in medicine. Erlbaum, Hillsdale, NJ

Munz Y ¹, Moorthy K, Bann S, Shah J, Ivanova S, Darzi SA.

Ceiling effect in technical skills of surgical residents.

Am J Surg. 2004 Sep;188(3):294-300

Munz Y ², Kumar BD, Moorthy K, Bann S, Darzi A.

Laparoscopic virtual reality and box trainers: is one superior to the other?

Surg Endosc. 2004 Mar;18(3):485-94.

National Confidential Enquiry into Perioperative Death 1988

www.ncepod.org.uk

Neale G, Woloshynowych M, Vincent C. Exploring the causes of adverse events in NHS hospital practice.

J R Soc Med. 2001;94:322-30.

Neary¹ W D, Heather B P, Earnshaw J J. The physiological and operative severity score for the enumeration of mortality and morbidity (POSSUM).

Br J Surg 2003; 90: 157-165.

Neary² WD, Crow P, Foy C, Prytherch D, Heather BP, Earnshaw JJ.

Comparison of POSSUM scoring and the Hardman Index in selection of patients for repair of ruptured abdominal aortic aneurysm.

Br J Surg. 2003 Apr;90(4):421-5.

Neumayer L, Giobbie-Hurder A, Jonasson O, Fitzgibbons R Jr, Dunlop D, Gibbs J, Reda D, Henderson W; Veterans Affairs Cooperative Studies Program 456 Investigators.

Open mesh versus laparoscopic mesh repair of inguinal hernia.

N Engl J Med. 2004 Apr 29;350(18):1819-27

Nielsen PE, Foglia LM, Mandel LS, Chow GE.

Objective structured assessment of technical skills for episiotomy repair.

Am J Obstet Gynecol. 2003 Nov;189(5):1257-60.

Qayumi AK, Cheifetz RE, Forward AD, Baird RM, Litherland HK, Koetting SE.

Teaching and evaluation of basic surgical techniques: the University of British Columbia experience.

J Invest Surg. 1999 Nov-Dec;12(6):341-50.

O'Toole RV, Playter RR, Krummel TM, Blank WC, Cornelius NH, Roberts WR, Bell WJ, Raibert M.

Measuring and developing suturing technique with a virtual reality surgical simulator.

J Am Coll Surg. 1999 Jul;189(1):114-27.

Paisley AM¹, Baldwin PJ, Paterson-Brown S.

Validity of surgical simulation for the assessment of operative skill.

Br J Surg. 2001 Nov;88(11):1525-32.

Paisley AM ², Baldwin P, Paterson-Brown S.

Feasibility, reliability and validity of a new assessment form for use with basic surgical trainees.

Am J Surg. 2001 Jul;182(1):24-9.

Porter GA, Soskolne CL, Yakimets WW, Newman SC.

Surgeon-related factors and outcome in rectal cancer.

Ann Surg 1998;227:157-67.

Powers TW, Murayama KM, Toyama M, Murphy S, Denham EW 3rd, Derossis AM, Joehl RJ.

Housestaff performance is improved by participation in a laparoscopic skills curriculum.

Am J Surg. 2002 Dec;184(6):626-9

Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ.

POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity.

Br J Surg. 1998 Sep;85(9):1217-20.

Putnis S, Merville-Tugg R, Atkinson S.

'One-stop' inguinal hernia surgery--day-case referral, diagnosis and treatment.

Ann R Coll Surg Engl. 2004 Nov;86(6):425-7

Quinn F, Keogh P, McDonald A, Hussey D.

A study comparing the effectiveness of conventional training and virtual reality simulation in the skills acquisition of junior dental students.

Eur J Dent Educ. 2003 Nov;7(4):164-9.

Rasmussen J, 1983. Skills, rules knowledge: signals, signs and symbols and other distinctions in human performance models. IEEE Trans. Systems, Man and Cybernetics, SMC-13: 257-267

Reason¹ J, 1990. Human Error, 1st edition
Cambridge University Press

Reason² J.
Understanding adverse events: human factors
Quality in Health Care 1995; 4:80-89

Regehr G, MacRae H, Reznick RK, Szalay D.
Comparing the psychometric properties of checklists and global rating scales for
assessing performance on an OSCE-format examination.
Acad Med. 1998 Sep;73(9):993-7.

Reznick R¹, Regehr G, MacRae H, Martin J, McCulloch W.
Testing technical skill via an innovative "bench station" examination.
Am J Surg. 1997 Mar;173(3):226-30.

Reznick R.K.²
Virtual reality surgical simulators: feasible but valid?
J Am Coll Surg. 1999 Jul;189(1):127-8.

Richards C, Rosen J, Hannaford B, Pellegrini C, Sinanan M.
Skills evaluation in minimally invasive surgery using force/torque signatures.
Surg Endosc. 2000 Sep;14(9):791-8.

Richardson MC, Bell G, Fullarton GM.
Incidence and nature of bile duct injuries following laparoscopic cholecystectomy: an
audit of 5913 cases. West of Scotland Laparoscopic Cholecystectomy Audit Group.
Br J Surg. 1996 Oct;83(10):1356-60.

Risucci DA.¹
Visual spatial perception and surgical competence.
Am J Surg. 2002 Sep;184(3):291-5.

Risucci DA², Tortolani AJ, Ward RJ.
Ratings of surgical residents by self, supervisors and peers.
Surg Gynecol Obstet. 1989 Dec;169(6):519-26.

Ritter EM, McClusky DA 3rd, Lederman AB, Gallagher AG, Smith CD.

Objective psychomotor skills assessment of experienced and novice flexible endoscopists with a virtual reality simulator.

J Gastrointest Surg. 2003 Nov;7(7):871-7

Rogers DA ¹, Regehr G, Yeh KA, Howdieshell TR.

Computer-assisted learning versus a lecture and feedback seminar for teaching a basic surgical technical skill.

Am J Surg. 1998 Jun;175(6):508-10.

Rogers DA ², Regehr G, Howdieshell TR, Yeh KA, Palm E.

The impact of external feedback on computer-assisted learning for surgical technical skill training.

Am J Surg. 2000 Apr;179(4):341-3.

Rogers DA ³, Regehr G, MacDonald J.

A role for error training in surgical technical skill instruction and evaluation.

Am J Surg. 2002 Mar;183(3):242-5.

Rosser JC, Rosser LE, Savalgi RS.

Skill acquisition and assessment for laparoscopic surgery.

Arch Surg. 1997 Feb;132(2):200-4.

Rosen J ¹, Solazzo M, Hannaford B, Sinanan M.

Objective laparoscopic skills assessments of surgical residents using Hidden Markov Models based on haptic information and tool/tissue interactions.

Stud Health Technol Inform. 2001;81:417-23.

Rosen J ², Solazzo M, Hannaford B, Sinanan M.

Task decomposition of laparoscopic surgery for objective evaluation of surgical residents' learning curve using hidden Markov model.

Comput Aided Surg. 2002;7(1):49-61.

Runciman WB, Sellen A, Webb RK, Williamson JA, Currie M, Morgan C.

Errors, incidents and accidents in anaesthetic practice.

Anaesthesia and Intensive Care 1993;21:506-18.

Russell JC, Walsh SJ, Mattie AS, Lynch JT.

Bile duct injuries, 1989-1993. A statewide experience. Connecticut Laparoscopic Cholecystectomy Registry.

Arch Surg. 1996 Apr;131(4):382-8.

Saifi J, Chang BB, Paty PS, Kaufman J, Leather RP, Shah DM.

An animal model for instructing and the study of in situ arterial bypass.

J Vasc Surg. 1990 Nov;12(5):538-40.

Sarker S.K.¹

Courses, counsellors & cadavers: reducing errors in the operating theatre

BMJ 2003; 327: s109

Sarker S.K.², Vincent C., Darzi A.

Assessing the Teaching of Technical Skills

Am J Surg 2005 April; 189(4):416-418

Sarker S.K.³ Chang A. Vincent C. Darzi A.W.

Dynamic Migration of Technical Errors in Performing Laparoscopic Cholecystectomy

Surg Endosc 2005 : 19 ; s11

Sandermann J, Panduro Jensen L.

The Danish specialist training in vascular surgery.

Eur J Vasc Endovasc Surg. 2002 Apr;23(4):353-7

Satava¹ RM, Cuschieri A, Hamdorf J.

Metrics for objective assessment.

Surg Endosc 2003;17: 220-6.

Satava² RM.

The nature of surgical error: a cautionary tale and a call to reason.

Surg Endosc. 2005 Aug;19(8):1014-6

Savader SJ, Lillemoe KD, Prescott CA, Winick AB, Venbrux AC, Lund GB, Mitchell SE, Cameron JL, Osterman FA Jr.

Laparoscopic cholecystectomy-related bile duct injuries: a health and financial disaster.

Ann Surg. 1997 Mar;225(3):268-73

Schwartz RW, Donnelly MB, Sloan DA, Johnson SB, Strodel WE.

The relationship between faculty ward evaluations, OSCE, and ABSITE as measures of surgical intern performance.

Am J Surg. 1995 Apr;169(4):414-7

Schell SR, Flynn TC.

Web-based minimally invasive surgery training: competency assessment in PGY 1-2 surgical residents.

Curr Surg. 2004 Jan-Feb;61(1):120-4.

Scott ¹ DJ, Valentine RJ, Bergen PC, Rege RV, Laycock R, Tesfay ST, Jones DB.

Evaluating surgical competency with the American Board of Surgery In-Training Examination, skill testing, and intraoperative assessment.

Surgery. 2000 Oct;128(4):613-22.

Scott ² DJ, Bergen PC, Rege RV, Laycock R, Tesfay ST, Valentine RJ, Euhus DM, Jeyarajah DR, Thompson WM, Jones DB.

Laparoscopic training on bench models: better and more cost effective than operating room experience?

J Am Coll Surg. 2000 Sep;191(3):272-83.

Senders J. & Moray N.

Human Error: Cause, Prediction, and Reduction.

1991 Lawrence Erlbaum Associates: Hillsdale

Seymour NE ¹, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM.

Virtual reality training improves operating room performance: results of a randomized, double-blinded study.

Ann Surg. 2002 Oct;236(4):458-63

Seymour NE ², Gallagher AG, Roman SA, O'Brien MK, Andersen DK, Satava RM.
Analysis of errors in laparoscopic surgical procedures.
Surg Endosc. 2004 Apr;18(4):592-5.

Sexton JB, Thomas EJ, Helmreich RL.
Error, stress and teamwork in medicine and
aviation: cross sectional surveys.
BMJ 2000;320:745-9.

Shah J, Buckley D, Frisby J, Darzi A.
Reaction time does not predict surgical skill.
Br J Surg. 2003 Oct;90(10):1285-6.

Shepherd A
Analysis and training in information technology tasks in:
Diaper D, Task Analysis for human computer interaction
1989 1st edition, Ellis Horwood

Sjoerdsma W, Herder JL, Horward MJ, Jansen A, Bannenberg JJG,
Grimbergen CA Force transmission of laparoscopic grasping instruments.
Min Invas Ther Allied Technol 1997 ; 6: 274–278

Smith ¹ S, Wan A, Taffinder N, Read S, Emery R, Darzi A.
Early experience and validation work with Procedicus VA--the Prosolvira virtual
reality shoulder arthroscopy trainer.
Stud Health Technol Inform.1999;62:337-43.

Smith ² SG, Torkington J, Brown TJ, Taffinder NJ, Darzi A.
Motion analysis.
Surg Endosc. 2002 Apr;16(4):640-5.

Smith CD, Farrell TM, McNatt SS, Metreveli RE.
Assessing laparoscopic manipulative skills.
Am J Surg. 2001 Jun;181(6):547-50.

Starkes JL, Payk I, Hodges NJ.

Developing a standardized test for the assessment of suturing skill in novice microsurgeons.

Microsurgery. 1998;18(1):19-22.

Steele RJ, Walder C, Herbert M.

Psychomotor testing and the ability to perform an anastomosis in junior surgical trainees.

Br J Surg. 1992 Oct;79(10):1065-7.

Stefanidis D, Korndorffer JR Jr, Sierra R, Touchard C, Dunne JB, Scott DJ.

Skill retention following proficiency-based laparoscopic simulator training.

Surgery. 2005 Aug;138(2):165-70

Summers AN, Rinehart GC, Simpson D, Redlich PN.

Acquisition of surgical skills: a randomized trial of didactic, videotape, and computer-based training.

Surgery. 1999 Aug;126(2):330-6.

Sun CC, Chiu AW, Chen KK, Chang LS.

Assessment of a three-dimensional operating system with skill tests in a pelvic trainer.

Urol Int. 2000;64(3):154-8.

Szalay D, MacRae H, Regehr G, Reznick R.

Using operative outcome to assess technical skill.

Am J Surg. 2000 Sep;180(3):234-7.

Taffinder N¹, Sutton C, Fishwick RJ, McManus IC, Darzi A.

Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator.

Stud Health Technol Inform. 1998;50:124-30.

Taffinder NJ², McManus IC, Gul Y, Russell RC, Darzi A.

Effect of sleep deprivation on surgeons' dexterity on laparoscopy simulator.

Lancet. 1998 Oct 10;352(9135):1191.

Taffinder N³, Smith S, Mair J, Russell R, Darzi A. Can a computer measure surgical precision? Reliability, validity and feasibility of the ICSAD.
Surg Endosc 1999;13(suppl 1): 81.

Tang¹ B, Hanna GB, Joice P, Cuschieri A.
Identification and categorization of technical errors by Observational Clinical Human Reliability Assessment (OCHRA) during laparoscopic cholecystectomy.
Arch Surg. 2004 Nov;139(11):1215-20.

Tang² B, Hanna GB, Cuschieri A.
Analysis of errors enacted by surgical trainees during skills training courses.
Surgery. 2005 Jul;138(1):14-20.
Tarcoveanu E, Niculescu D, Georgescu S, Epure O, Bradea C.
Conversion in laparoscopic cholecystectomy]
Chirurgia (Bucur). 2005 Sep-Oct;100(5):437-44.

Tekkis PP, McCulloch P, Poloniecki JD, Prytherch DR, Kessaris N, Steger AC.
Risk-adjusted prediction of operative mortality in oesophagogastric surgery with O-POSSUM.
Br J Surg. 2004 Mar;91(3):288-95.

Testoni PA, Sultan S, Baillie J.
Can computer simulation accelerate the development of procedural competence?
Am J Gastroenterol. 2004 Jan;99(1):38-9
Thomas¹ EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al.
Incidence and types of adverse events and negligent care in Utah and Colorado.
Med Care 2000; 38:261-271

Thomas² EJ, Brennan TA. Errors and adverse events in medicine: an overview.
In Vincent CA, Clinical risk management. Enhancing patient safety. London,
BMJ Publications, 2001.

Torkington J¹, Smith SG, Rees BI, Darzi A.
Skill transfer from virtual reality to a real laparoscopic task.
Surg Endosc. 2001 Oct;15(10):1076-9.

Torkington J ², Smith SG, Rees B, Darzi A.

The role of the basic surgical skills course in the acquisition and retention of laparoscopic skill.

Surg Endosc. 2001 Oct;15(10):1071-5.

Tsumura H, Ichikawa T, Hiyama E, Murakami Y.

Laparoscopic and open approach in perforated peptic ulcer.

Hepatogastroenterology. 2004 Sep-Oct;51(59):1536-9.

Van Rij ¹ AM, Jiang P, Solomon C, Christie RA, Hill GB. Recurrence after varicose vein surgery: a prospective long-term clinical study with duplex ultrasound scanning and air plethysmography.

J Vasc Surg. 2003 Nov;38(5):935-43

Van Rij ² AM, Chai J, Hill GB, Christie RA. Incidence of deep vein thrombosis after varicose vein surgery.

Br J Surg. 2004 Dec;91(12):1582-5

Van Sickle KR, McClusky DA 3rd, Gallagher AG, Smith CD.

Construct validation of the ProMIS simulator using a novel laparoscopic suturing task.

Surg Endosc. 2005 Sep;19(9):1227-31

Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondre K, Stanbridge D, Fried GM.

A global assessment tool for evaluation of intraoperative laparoscopic skills.

Am J Surg. 2005 Jul;190(1):107-13

Velmahos GC, Toutouzas KG, Sillin LF, Chan L, Clark RE, Theodorou D, Maupin F. Cognitive task analysis for teaching technical skills in an inanimate surgical skills laboratory.

Am J Surg. 2004 Jan;187(1):114-9.

Vincent ¹ C, Neale G, Woloshynowych M.

Adverse events in British hospitals: preliminary retrospective record review.

BMJ 2001;322:517-9.

Vincent² C.A.

Patient Safety

2006 London, Elsevier Churchill Livingstone

Vogt VY, Givens VM, Keathley CA, Lipscomb GH, Summitt RL Jr.

Is a resident's score on a videotaped objective structured assessment of technical skills affected by revealing the resident's identity?

Am J Obstet Gynecol. 2003 Sep;189(3):688-91.

Wanzel KR¹, Matsumoto ED, Hamstra SJ, Anastakis DJ.

Teaching technical skills: training on a simple, inexpensive, and portable model.

Plast Reconstr Surg. 2002 Jan;109(1):258-63.

Wanzel² KR, Hamstra SJ, Anastakis DJ, Matsumoto ED, Cusimano MD.

Effect of visual-spatial ability on learning of spatially-complex surgical skills.

Lancet. 2002 Jan 19;359(9302):230-1.

Wanzel KR³, Hamstra SJ, Caminiti MF, Anastakis DJ, Grober ED, Reznick RK.

Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance.

Surgery. 2003 Nov;134(5):750-7.

Ward M, MacRae H, Schlachta C, Mamazza J, Poulin E, Reznick R, Regehr G.

Resident self-assessment of operative performance.

Am J Surg. 2003 Jun;185(6):521-4

Watkins JL, Etzkorn KP, Wiley TE, DeGuzman L, Harig JM.

Assessment of technical competence during ERCP training.

Gastrointest Endosc. 1996 Oct;44(4):411-5.

Watterson JD, Beiko DT, Kuan JK, Denstedt JD.

Randomized prospective blinded study validating acquisition of ureteroscopy skills using computer based virtual reality endourological simulator.

J Urol. 2002 Nov;168(5):1928-32.

Wetzel CM, Kneebone RL, Woloshynowych M, Nestel D, Moorthy K, Kidd J, Darzi A.

The effects of stress on surgical performance.

Am J Surg. 2006 Jan;191(1):5-10

Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study.

Med J Aust. 1995 Nov 6;163(9):458-71.

Winckel CP, Reznick RK, Cohen R, Taylor B.

Reliability and construct validity of a structured technical skills assessment form.

Am J Surg. 1994 Apr;167(4):423-7.

Windsor JA, Zoha F.

The laparoscopic performance of novice surgical trainees: testing for acquisition, loss, and reacquisition of psychomotor skills.

Surg Endosc. 2005 Aug;19(8):1058-63

Winterborn RJ, Foy C, Earnshaw JJ. Causes of varicose vein recurrence: late results of a randomized controlled trial of stripping the long saphenous vein.

J Vasc Surg. 2004 Oct;40(4):634-9

Wong K, Stewart F.

Competency-based training of basic trainees using human cadavers.

ANZ J Surg. 2004 Aug;74(8):639-42.

www.bristol-inquiry.org.uk

www.drfooster.co.uk

www.jchst.org

Publications & Presentations

Papers & Articles

1. Assessing Technical & Technological Skills in Laparoscopic Surgery
Sarker S.K., Chang A., Vincent C.
Journal of the Society of Laparoendoscopic Surgeons 2006; 10 (3): 284-92
2. Self-Appraisal Hierarchical Task Analysis of Expert Surgeons Performing Laparoscopic Surgery
Sarker S.K. Hutchinson R. Chang A. Vincent C. Darzi A.W.
Surgical Endoscopy 2006 Apr;20(4):636-40
3. Development of Assessing Generic & Specific Technical Skills in Laparoscopic Surgery
Sarker S.K., Chang A., Vincent C. Darzi A.
American Journal of Surgery 2006 Feb;191(2):238-44.
4. Errors in Surgery
Sarker S.K. , Vincent C.
International Journal of Surgery 2005; 3(1):75-81
5. Technical Skills Errors in Laparoscopic Cholecystectomies by Expert Surgeons
Sarker S.K., Chang A., Vincent C. Darzi A.
Surgical Endoscopy 2005; 19(4):832-835
6. Assessing the Teaching of Technical Skills
Sarker S.K. , Vincent C., Darzi A.
American Journal of Surgery 2005; 189(4):416-418
7. Systems Approaches to Surgical Quality & Safety:
From concept to measurement
Vincent C., Moorthy K., **Sarker S.K.**, Chang A., Darzi A.
Annals of Surgery 2004; 239(4):475-82
8. Operation Surgical Safety
Sarker S.K.
Bulletin of the Royal College of Physicians & Surgeons of Glasgow 2004; 33 (3) 2
9. Objective Assessment of Technical Skills in Surgery
Moorthy K., Muntz Y., **Sarker S.K.**, Darzi A.
British Medical Journal 2003; 327:1032-7
10. Courses, Counsellors & Cadavers: reducing errors in the operating theatre
Sarker S.K.
British Medical Journal 2003; 327: s109

Presentations to Societies & Abstract Publications

1. Self-Appraisal of Senior Surgeons Using Objective Structured Assessment of Technical Skill in Laparoscopic Cholecystectomies
Sarker S.K. , Chang A., Vincent C. , Darzi A.
Society of American Gastrointestinal Endoscopic Surgeons, Denver, USA,
April 2004
Surgical Endoscopy 2004 : 18 ; s233
2. Surgical Technical Errors in Performing Open & Laparoscopic Surgery
Sarker S.K. Chang A. Vincent C. Darzi A.W.
Association of Surgeons of Great Britain & Ireland, Harrogate, April 2004
British Journal of Surgery 2004: 91; s78
3. Dynamic Migration of Technical Errors in Performing Laparoscopic Cholecystectomy
Sarker S.K. Chang A. Vincent C. Darzi A.W.
European Association of Endoscopic Surgeons, Barcelona, June 2004
Surgical Endoscopy 2005 : 19 ; s11
4. Technical Styles in Performing Laparoscopic Cholecystectomies
Sarker S.K. Hutchinson R. Chang A. Vincent C. Darzi A.W.
Association of Upper-GI Surgeons of Great Britain & Ireland, Cambridge,
September 2004
5. Development of a Structured Global Technical Skills Assessment Tool in Open & Laparoscopic Surgery
Sarker S.K. , Chang A., Vincent C. , Darzi A.
Society of American Gastrointestinal Endoscopic Surgeons, Hollywood, USA,
April 2005
Surgical Endoscopy 2005 : 19 ; s248
6. Self-Appraisal of Consultant Surgeons Performing Laparoscopic Cholecystectomies Using Hierarchical Task Analysis
Sarker S.K. , Hutchinson R. , Chang A., Vincent C. , Darzi A.
Society of American Gastrointestinal Endoscopic Surgeons, Hollywood, USA,
April 2005
Surgical Endoscopy 2005 : 19 ; s248
7. Construct Validity of Assessors of Structured Surgical Technical Skills Assessment in Laparoscopic Surgery
Sarker S.K. , Chang A., Vincent C. , Darzi A.
Society of American Gastrointestinal Endoscopic Surgeons, Hollywood, USA,
April 2005
Surgical Endoscopy 2005 : 19 ; s248

Presentations to Societies & Abstract Publications

8. A Reliability Analysis of Video-Based Rating Scales for Technical Skills Assessments in Laparoscopic Surgery
Aggarwal R., Grantcharov T., Moorthy K., Papasavas P., Milland T.,
Sarker S.K., Darzi A.
Society of American Gastrointestinal Endoscopic Surgeons,
Hollywood, USA, April 2005
Surgical Endoscopy 2005 : 19 ; s286
9. Assessment of Technical Skills in Open & Laparoscopic Surgery
Sarker S.K. Chang A. Vincent C. Darzi A.W.
Association of Surgeons of Great Britain & Ireland, Glasgow, April 2005
British Journal of Surgery 2005: 92; s22-23
10. Training in the Assessment of Technical Skills in Laparoscopic Surgery
Sarker S.K. , Chang A., Vincent C. , Darzi A.
European Association of Endoscopic Surgeons, Venice, Italy, June 2005
Surgical Endoscopy 2006
12. Assessment of Technical & Technological Skills in Laparoscopic Surgery
Sarker S.K. , Chang A., Vincent C. , Darzi A.
Society of Laparoendoscopic Surgeons, San Diego, USA, September 2005
Journal of the Society of Laparoendoscopic Surgeons 2005: 9 (3) ; s33
13. Generic & Specific Technical Skills Errors by Consultant Surgeons in Elective Open Surgery
Sarker S.K. , Chang A., Vincent C.
Association of Surgeons of Great Britain & Ireland, Edinburgh, May 2006
14. Migration of Technical Skills Errors by Expert Surgeons in Laparoscopic Surgery
Sarker S.K. , Chang A., Vincent C.
Association of Surgeons of Great Britain & Ireland, Edinburgh, May 2006
15. Assessing the Teaching of Technical Skills in Operations & Procedures
Sarker S.K. Amygdalos I. Darzi A. Kneebone R.
Association of Surgeons of Great Britain & Ireland, Edinburgh, May 2006
British Journal of Surgery 2006; 10 (3): 284-292

Appendix

Appendix Chapter 1 - Surgical Outcome

Table 1.1 – Measures of Surgical Outcome

Measures of Surgical Outcome
Pain
Length of hospital stay
Return to work by patient
Mobility of patient
Self-caring
Completion of the operation
Recurrence of condition e.g. hernia
Recurrence of disease e.g. cancer
Post-operative complications – Morbidity
Post-operative death – Mortality

Table 1.2 – Types of Operation Affecting Outcome

Type of Operation Affecting Outcome
Elective or Emergency
Open or Closed
Dirty or Clean
Operative Severity e.g. major
Type of Surgery e.g. orthopaedics

Table 1.3 – Patient Factors Affecting Surgical Outcome

Patient Factors	
Age	Electro Cardiograph findings
Sex	Chest X-ray
BMI (Obesity index)	White cell count
ASA Grade	Anatomical site of condition/disease
Glasgow Coma Scale	Recurrence of disease or condition
Cardiac Failure	Diagnostic sub-group (Benign or Malignant)
Respiratory Disease	Size of tumour
Blood Pressure	Tumour progression
Pulse	Type of tumour
Haemoglobin level	Previous surgery to patient
Urea level	Previous adjuvant therapy to patient
Electrolytes	Disease complications

Table 1.4 – Non-Patient Factors Affecting Outcome

Non-Patient Factors	
Time of operation (Night or Day)	Type of nursing care post-operatively
Surgeon's experience	Nurse's experience
Length of operation	Nurse's technique
Surgeon's stresses e.g. amount of sleep	Theatre Staff's experience
Surgeon's technique	Theatre Staff's technique
Blood loss	Doctor's post-operative experience
Anaesthetist's experience	Doctor's post-operative techniques
Anaesthetist's technique	

Table 1.5 – Organisation & Equipment Affecting Surgical Outcome

Organisation & Equipment Factors
Equipment/instrument unavailable
Faulty equipment
Transfer of patient to another hospital
Availability of surgical bed
Availability of HDU/ICU bed

Table 1.6 – APACHE II Score (Acute Physiology Score) + (Age Points) + (Chronic Health Points)

Acute Physiology Score	Chronic Health Points
1 = Rectal temp (C)	Non-operative patients = 5
2 = Mean arterial pressure (mmHg)	Emergency postoperative patients = 5
3 = Heart rate (bpm)	Elective postoperative patients = 2
4 = Respiratory rate (bpm)	
5 = Oxygen delivery (ml/min)	
6 = PO ₂ (mmHg)	
7 = arterial pH	
8 = Serum sodium (mmol/l)	
9 = Serum potassium (mmol/l)	
10 = Serum creatinine (mg/dl)	
11 = Haematocrit (%)	
12 = White cell count (10 ³ /ml)	

Table 1.7 – POSSUM Scoring

Physiological assessment Provides exponential score on 12 variables	Operative severity Provides exponential score on 6 variables
Age	Operative magnitude
Cardiac signs	Number of operations within 30 days
Respiratory signs	Blood loss
Systolic blood pressure	Peritoneal contamination
Pulse	Presence of malignancy
Coma score	Timing of operation
Serum urea	
Serum sodium	
Serum potassium	
Haemoglobin	
White cell count	
ECG	

Table 1.8 – Framework of Factors Influencing Clinical Practice

FACTOR TYPES	INFLUENCING CONTRIBUTORY FACTORS
Institutional Context	Economic and regulatory context National health service executive Clinical negligence scheme for trusts
Organisational and Management Factors	Financial resources & constraints Organisational structure Policy standards and goals Safety culture and priorities
Work Environment Factors	Staffing levels and skills mix Workload and shift patterns Design, availability and maintenance of equipment Administrative and managerial support
Team Factors	Verbal communication Written communication Supervision and seeking help Team structure (consistency, leadership, etc)
Individual (staff) Factors	Knowledge and skills Competence Physical and mental health
Task Factors	Task design and clarity of structure Availability and use of protocols Availability and accuracy of test results
Patient Factors	Condition (complexity & seriousness) Language and communication Personality and social factors

Appendix Chapter 2 - Assessment of Technical Skills in Surgery

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
130	2005	Moorthy ³ et al	London UK	27	Novice & Intermediate	Bench Model Global Scoring Open	Reliability good, construct validity demonstrated	Simulated environment could be used to assess technical & team skills
129	2005	Korndorffer et al	New Orleans USA	17	Novice & Intermediate	Real Animal Bench Model Error Scoring MIS	Improvement in training group	Training on bench model improves operative performance
128	2005	Stefanidis et al	New Orleans USA	14	Novice	Bench Model Virtual Reality Motion Analysis, Time & Error Scoring MIS	Skill retention better for bench model than virtual reality	May be useful in skills training
127	2005	Dosis ² et al	London UK	5	Expert & Intermediate	Real Human Motion Analysis MIS	Difference between motion analysis between two groups	Can be used to assess operative procedures
126	2005	Eriksen & Grantcharov	Copenhagen Denmark	24	Expert & Intermediate	Virtual Reality Motion Analysis, Time & Error Scoring MIS	Significant differences between two groups	Can be used in surgical training

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
125	2005	Windsor et al	Auckland New Zealand	10	Novice	Virtual Reality Motion Analysis, Time & Error Scoring MIS	Good improvement in technical skills by novices	Can assess the acquisition of laparoscopic skills
124	2005	Vassiliou et al	Montreal Canada	21	Intermediate & Expert	Real Human Global & Checklist Scoring MIS	Good reliability & construct validity	Good in assessing laparoscopic skill
123	2005	Hance ¹ et al	London UK	40	Intermediate & Expert	Bench Model Global Scoring Open	Good reliability & construct validity	Construct validity in bench models for cardiac surgeons
122	2005	Backstein ² et al	Toronto Canada	26	Intermediate	Bench Model Global & Checklist Score Open	No effect of video feedback	No improvement of surgical performance using video feedback
121	2005	Van Sickle et al	Atlanta USA	10	Novice & Expert	Virtual Reality Motion Analysis MIS	Difference with experts & novices	Construct validity
120	2005	Beard ² et al	Sheffield UK	74	Expert	Bench Model Global Scoring Open	Most participants had 75-100% score	May have criterion validity

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
119	2005	Gosman et al	Pittsburgh USA	31	Intermediate & Expert	Real Human Global Scoring MIS	Faculty members completed forms in 36% cases	Expert compliance in completing assessments was low
118	2005	Hance ² et al	London UK	37	Intermediate	Bench Model Motion Analysis MIS	Improvement in motion analysis by participants	Can assess psychomotor skills during courses
117	2005	Larson et al	Illinois USA	N/K	Intermediate	Real Human Global Scoring Open MIS	Inter-rater reliability good	Useful in tracking residents progress
116	2005	Bann ⁶ et al	London UK	11	Novice & Intermediate	Bench Model Virtual Reality Motion Analysis, Time & Error Scoring Open MIS	Good reliability Inter-rater $r = 0.77-0.94$	Assessment is reliable
115	2005	McClusky et al	Atlanta USA	11	Novice	Virtual Reality Motion Analysis & Error Scoring MIS	Good reliability Perceptual ability $r = -0.76$, psychomotor skills $r = 0.62$	Number of practices to achieve performance goals on VR dependent on perceptual & psychomotor aptitude

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
114	2005	Arora et al	New York USA	38	Novice & Intermediate	Virtual Reality Motion Analysis & Error Scoring MIS	Good Correlation between VR simulators $r = 0.43 - 0.63$	Construct validity of new endoscopic sinus simulator
113	2005	Goff ⁵ et al	Portland USA	116	Intermediate	Bench Models Virtual Reality Global Scoring & Checklists MIS	Good reliability & construct validity	Large scale testing confirms OSATS as reliable & valid for assessing surgical skills
112	2005	Beard ¹ et al	Sheffield UK	33	Intermediate	Bench Models Checklist Open	Inter-rater reliability $r = 0.96$	Feasible & reliable to assess technical skills by direct & video recordings
111	2005	Adamsen et al	Copenhagen Denmark	24	Novice & Experts	Virtual Reality Motion Analysis, Time & Error Scoring MIS	Construct validity between groups	Laparoscopic & endoscopic simulators able to detect grades of operators
110	2005	Bann ⁵ et al	London UK	38	Intermediate	Bench Models Global & Error Scoring Open	Correlation between error & global scores	May be useful in reducing surgical learning curve
109	2005	Katz et al	Creteil France	44	Intermediate & Expert	Bench Models Time MIS	Construct validity	Possible criterion level for operating rights

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
108	2004	Fried et al	Montreal Canada	200	Novice, Intermediate & Expert	Bench Models Global Scoring MIS	Face, construct, concurrent validities	Practical & inexpensive system to teach & assess laparoscopic skills
107	2004	Grober ³ et al	Toronto Canada	50	Novice	Bench Models Check List & Global Open	Hands on training improves technical skill	Training on low fidelity as good as high fidelity models
106	2004	Munz ¹ et al	London UK	30	Novice & Intermediate	Bench Models Motion analysis & Global Open	No construct validity	Ceiling effect of generic skills in trainees
105	2004	Wong et al	Sydney Australia	8	Novice	Bench Models Time & Error Scoring Open	Time and errors committed reduced after training	Anatomy courses improves confidence & technical skills
104	2004	Seymour ² et al	New Haven USA	14	Intermediate	Real Human Error Scoring MIS	Inter-rater reliability 84-100%	Procedural errors had good interrater agreement
103	2004	Gallagher ⁵ et al	Atlanta USA	124	Novice, intermediate & Expert	Virtual Reality Motion analysis & Error scoring MIS	With practice novices improved	VR can discriminate between experts & novices

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
102	2004	Kohls-Gatzoulis et al	Toronto Canada	21	Novice	Bench Models Global & Didactic Open	Didactic knowledge same in each group, but cognitive group better at error detection. Technical skills improved in both groups	Training improves generic & specific technical skills
101	2004	Tang ¹ et al	Dundee UK	26	Intermediate & Expert	Real Human Error Scoring MIS	Technical errors dependent on technique used	Error scoring system able to document technical errors
100	2004	Datta ⁵ et al	London UK	22	Novice, Intermediate & Expert	Real Human & Bench Models Check Lists & Global Open	Correlation of surgeon's grade to performance on bench model & real operation	Performance on bench model comparable to real operation
99	2004	Grober ² et al	Toronto Canada	18	Novice	Bench Models Check Lists & Global Open	Retention of technical skills in bench model group	Training can improve novice surgeons
98	2004	Crawford & Colt	San Diego USA	12	Novice	Virtual Reality Check List MIS	No relation between technical skill, knowledge and experience	Practice on VR may improve technical skills

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
97	2004	Backstein ¹ et al	Toronto Canada	29	Novice & Intermediate	Bench Models Check Lists & Global Open	No difference between groups	Feedback did not improve technical skills
96	2004	Eversbusch et al	Copenhagen Denmark	28	Novice & Experts	Virtual Reality Motion Analysis, Time & Error Scoring MIS	Psychomotor training improved performance	Different learning curves between grades, training improved performance
95	2004	Velmahos et al	Los Angeles USA	26	Novice	Bench Models Check Lists Open	Improvement in technical skills in course group	Surgical course improves technical skills and knowledge
94	2004	Bann ⁴ et al	London UK	30	Novice & Intermediate	Bench Models Motion Analysis & Error Scoring Open	Intermediates better in technical skill and equipment knowledge but not in error analysis	Surgical error analysis should be more formally taught
93	2004	Feldman et al	Montreal Canada	50	Novice	Bench Models Time & Error Scoring MIS	High scores on MISTELS correlated to high scores in the training assessment	In training assessment did not identify residents with below average technical skill

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
92	2004	Grantcharov ⁴ et al	Copenhagen Denmark	16	Intermediate	Real Human Virtual Reality Motion Analysis, Time & Error Scoring MIS	VR training improved performance	VR training improved performance in real operations
91	2004	Schell et al	Gainesville USA	22	Novice	Bench Models Didactic MIS	Training group were better in laparoscopic skills & knowledge	Training should be implemented in year 2 residents
90	2004	Munz ² et al	London UK	24	Novice	Virtual Reality Motion analysis & Error Scoring MIS	Trained groups were better at technical skills	Training improves laparoscopic skills
89	2004	Hernandez et al	London USA	13	Intermediate & Expert	Bench Models Motion analysis & Global MIS	Both groups of surgeons improved over time	Using the robotic system inexperienced surgeons became competent in a complex laparoscopic task
88	2003	Grantcharov ³ et al	Copenhagen Denmark	41	Novice, Intermediate & Expert	Virtual Reality Motion Analysis, Time & Error Scoring MIS	Construct validity between groups	Different learning curves for different grades

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
87	2003	Moorthy ¹ et al	London UK	18	Novice & Expert	Virtual Reality Check Lists & Error Scoring MIS	Improvement of technical skills in junior group	VR bronchoscopy is an effective training tool
86	2003	Wanzel ³ et al	Toronto Canada	47	Novice, Intermediate & Expert	Bench Models Motion analysis, Global & Check List Open	Significant difference between groups in surgical performance	These tests not recommended for selection of residents
85	2003	Nielsen et al	Seattle USA	18	Novice	Bench Models Global & Check Lists Open	Good inter-rater reliability & validity	Reliable method of assessing resident technical skills
84	2003	Mackay ³ et al	London UK	21	Novice & Experts	Bench Models Virtual Reality Motion analysis, Global, Check List, Error Scoring & Time Open MIS	Good inter-rater reliability & validity	Feasible & effective method of assessing technical skills in basic surgical trainees
83	2003	Adrales et al	Lexington USA	27	Novice, intermediate & Experts	Bench Models Global & Check Lists MIS	Face & construct validity	Quality of performance increased with experience

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
82	2003	Vogt et al	Memphis USA	33	Intermediate	Bench Models Global & Check Lists Open	Different between the two views and reviewers	Bias can occur in evaluation of videotaped technical skills
81	2003	Ritter et al	Atlanta USA	13	Novice & Intermediate	Virtual Reality Motion analysis & Error scoring MIS	Good reliability & validity, improvement of novices with practice	Can distinguish between novice & intermediate
80	2003	Bann ¹ et al	London UK	70	Novice & Intermediate	Bench Models Virtual Reality Motion Analysis, Global, Time & Error Scoring Open MIS	Construct validity	Assessment of technical skills in trainees should be promoted and tracked
79	2003	Gettman et al	Dallas USA	20	Novice	Bench Models Global MIS	Predicted agreement with actual operative performance	Prediction of laparoscopic performance can be done
78	2003	Gallagher ³ et al	Seattle USA	195	Intermediate & Expert	Virtual Reality Bench Model Time, Motion Analysis & Error Scoring MIS	Variability in intermediate group in performance	Assessment of laparoscopic skills is possible

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
77	2003	Eastridge et al	Dallas USA	35	Intermediate	Virtual Reality Time, Motion Analysis & Error Scoring MIS	Time and number of errors increased after post on-call	Sleep deprivation & fatigue associated with increased errors
76	2003	Anastakis ¹ et al	Toronto Canada	50	Intermediate	Bench Models Global & Check List Open MIS	No difference between treatment and control groups	Training did not improve surgical skills
75	2003	Khan et al	London UK	89	Novice, Intermediate & Expert	Bench Models Motion Analysis & Global Open	Face, construct & concurrent validities	Can assess fundamentals of surgical technique
74	2003	Bann ² et al	London UK	89	Novice & Intermediate	Bench Models Error Scoring Open	Construct validity	Error analysis able to discriminate between grades
73	2003	Mahood et al	London UK	25	Novice, Intermediate & Expert	Virtual Reality Error scoring MIS	Construct validity	Difference in grade of operator to performance
72	2003	Quinn et al	Dublin Ireland	20	Novice	Virtual Reality Global Open	Difference between conventional and VR training groups	VR is unsuitable only method of feedback & evaluation in novices

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
71	2003	Gallagher ⁴ et al	Belfast UK	96	Novice & Experts	Virtual Reality Motion analysis & Error Scoring MIS	Good reliability & validity	VR can reliability predict performance
70	2003	Grober ¹ et al	Toronto Canada	50	Novice	Bench Models Motion Analysis & Global Open	Concurrent & construct validity for motion analysis	Motion analysis sensitive instrument for microsurgical performance
69	2003	Shah et al	London UK	41	Novice & Intermediate	Virtual Realty Time, Error & Motion Analysis MIS	No difference in time between grades	Time is not a measure of quality of performance
68	2003	Bann ³ et al	London UK	35	Novice, Intermediate & Expert	Bench Models Motion analysis & Time Open	Face & construct validity	Can assess dexterity
67	2003	Ward et al	Toronto Canada	26	Intermediate	Real Animal Global & Check List MIS	Correlation between experts & intermediates was moderate	Self-observation improved the intermediates to self-evaluate
66	2002	MacDonald et al	Illinois USA	21	Novice	Bench Models Time & Error Scoring Open	Time & errors made improved with time	Practice improves technical skills
65	2002	Mackay ¹ et al	London UK	4	Expert	Bench Models Motion Analysis & Time Open	Difference between fast and precise strategies	Time & movement is not fixed

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
64	2002	Datta ² et al	London UK	50	Novice, Intermediate & Expert	Bench Model Motion Analysis, Time & Global Open	Number of movements, time and global score discriminated between performance & grade	Correlation between motion analysis & global scoring
63	2002	Risucci ¹	New York USA	301	Novice, Intermediate & Expert	Bench Model Check List Open & MIS	High level visual spatial perception greater in experts	VSP proficiency a valid component to surgical competence
62	2002	Rogers ³ et al	Illinois USA	30	Novice	Bench Model Global & Error Scoring Open	Instruction enhanced acquisition of technical skill	No support for error training in improving surgical evaluation
61	2002	Powers et al	Chicago USA	27	Novice & Intermediate	Bench Model Error Scoring MIS	Improvement in novices & intermediates No construct validity	A curriculum improves technical skills in residents
60	2002	Wanzel ¹ et al	Toronto Canada	37	Novice	Bench Models Global & Check List Open	Improvement in after practice & feedback	Practice & feedback improves performance

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
59	2002	Coleman & Muller	Dallas USA	26	Intermediate	Bench Model Global MIS	Improvement in technical skills over time	Core curriculum improves technical performance
58	2002	Watterson et al	London Canada	20	Novice	Virtual Reality Global MIS	Post training improved global scores	VR training improves ureteroscopic skills
57	2002	Smith ² et al	London UK	15	Novice, Intermediate & Expert	Bench Models Motion Analysis & Time MIS	Construct validity	Motion analysis provides assessment laparoscopic dexterity
56	2002	Rosen ² et al	Seattle USA	12	Novice, Intermediate & Expert	Real Animal Mechanical & Motion Analysis MIS	Difference between experts & novices	Tool to assess laparoscopic skills
55	2002	Datta ³ et al	London UK	45	Novice, Intermediate & Expert	Virtual Reality Motion analysis, Time, Error scoring MIS	Construct validity	VR is a valid discriminator
54	2002	Mackay ² et al	London UK	41	Novice	Virtual Reality Time, Error Scoring & Motion Analysis MIS	Training improves laparoscopic skills	Benefit of practice in learning laparoscopic skills

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
53	2002	Seymour ¹ et al	Connecticut USA	16	Novice & Intermediate	Virtual Reality Time, Error & Motion Analysis MIS	Difference between trained & non- trained group	VR training improves laparoscopic skills
52	2002	Goff ⁴ et al	Seattle USA	16	Intermediate	Bench Models Global & Check Lists Open MIS	Inter-rater reliability & Construct Validity	Global scoring can assess intermediate technical skills
51	2002	Datta ⁴ et al	London UK	6	Intermediate	Bench Models Motion Analysis & Quality of Product Open	Manual dexterity correlated to outcome	Outcome of a procedure can be predicted by measuring surgical skill
50	2002	Ahlberg et al	Stockholm Sweden	29	Novice	Virtual Reality Bench Model Global, Error Scoring, Time & Motion Analysis MIS	No difference between groups	VR training did not improve surgical skill but predicted outcome

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
49	2001	Torkington ¹ et al	London UK	30	Novices	Virtual Reality Bench Model Motion Analysis MIS	Significant improvement in VR trained group	VR training transfers skill to a simple real task in novices
48	2001	Hamilton et al	Dallas USA	21	Intermediate	Bench Model Global MIS	Improvement in trained group	Curriculum improves knowledge & performance
47	2001	Friedlich et al	Toronto Canada	47	Novice & Intermediate	Bench Model Global & Check List Open	Difference between grades	Construct validity in assessing minor surgical skills
46	2001	McNatt & Smith	Atlanta USA	8	Novice & Expert	Virtual Reality Error, Time & Motion Analysis MIS	Experts performed better then novices	Can make objective assessment of acquisition of laparoscopic skill
45	2001	Torkington ² et al	London UK	13	Novice	Virtual Reality Error, Time & Motion Analysis MIS	Parameters improved after course	Course improves laparoscopic skill
44	2001	Datta ¹ et al	London UK	51	Novice, Intermediate & Experts	Bench Models Motion Analysis Open	Difference between grades	Motion analysis may be effective in measuring open surgical simulation

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
43	2001	Paisley ¹ et al	Edinburgh UK	89	Novice, Intermediate & Experts	Bench Models Virtual Reality Error Scoring, Time , Motion analysis & Check List Open MIS	No difference between groups	Assessment of technical skills need improvement
42	2001	Ault et al	Chicago USA	77	Novice & Intermediate	Bench Models Global & Check Lists Open	Good inter-rater reliability & construct validity	Assessment tool portable & feasible
41	2001	Goff ³ et al	Seattle USA	24	Novice & Intermediate	Bench Models Global & Check Lists Open	Good inter-rater reliability & construct validity	Useful tool to identify residents who need additional help
40	2001	Grantcharov ¹ et al	Copenhagen Denmark	14	Novice & Intermediate	Virtual Reality Motion Analysis, Time & Error Scoring MIS	Increased error rate, time & movements after night on call	Impaired speed and accuracy after night on call
39	2001	Grantcharov ² et al	Copenhagen Denmark	14	Novice & Intermediate	Virtual Reality Real Animals Motion Analysis, Time & Error Scoring MIS	Error scores & economy of movement correlated in VR & live animals	Correlation of performance on VR and live animal

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
38	2001	Smith et al	Atlanta USA	10	Novice	Bench Models Motion Analysis MIS	Repetition improved skills	Accuracy more important than time
37	2001	Rosen ¹ et al	Washington USA	8	Novice, Intermediate & Experts	Bench Models Motion Analysis MIS	Difference between groups	Can assess different levels of skill
36	2001	Paisley ² et al	Edinburgh UK	36	Novice	Real Human Check List Open MIS	Good reliability in Interaction, Knowledge, Teamwork, Technical Skills, but not for Clinical Skills	Feasible, reliable & valid for novices
35	2001	Francis ¹ et al	Dundee UK	40	Novice & Experts	Bench Models Motion Analysis & Error Scoring MIS	Difference between experts & novices	Eye-hand & dexterity better in experts but lower visuo-spatial ability
34	2000	MacRae et al	Toronto Canada	24	Intermediate	Bench Models Global Open MIS	Difference between intermediate grades	Reliable & valid test

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
33	2000	Goff ² et al	Seattle USA	24	Intermediate	Bench Models Global & Check Lists Open MIS	Good inter-rater reliability & construct validity	Assessment may identify those who need additional training
32	2000	Sun et al	Taipei China	34	Novice	Bench Model Time MIS	Decrease in performance time after 2 nd attempt	Practice improves laparoscopic skills
31	2000	Scott ² et al	Dallas USA	27	Novice	Real Human Bench Models Global scoring MIS	Difference between untrained & trained groups	Training improves skills
30	2000	Richards et al	Seattle USA	10	Novice & Experts	Bench Models Motion Analysis & Time MIS	Difference between experts & novices	Maybe helpful in training & assessment
29	2000	Rogers ² et al	Augusta USA	105	Novice	Bench Model Time & Check List Open	Feedback improved performance	Feedback & training improves performance
28	2000	Scott ¹ et al	Dallas USA	33	Novice	Bench Models Global MIS	No difference between novices	Programs should have multiple instruments of assessment

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
29	2000	Rogers ² et al	Augusta USA	105	Novice	Bench Model Time & Check List Open	Feedback improved performance	Feedback & training improves performance
28	2000	Scott ¹ et al	Dallas USA	33	Novice	Bench Models Global MIS	No difference between novices	Programs should have multiple instruments of assessment
27	2000	Szalay et al	Toronto Canada	20	Intermediate	Bench Models Global & Time Open	Difference between surgeons, moderate inter- rater reliability	Time & analysis of end product alternatives for assessment
26	1999	Eubanks et al	Seattle USA	30	Novice, Intermediate & Experts	Real Human Error Scoring MIS	Good inter-rater reliability	Technical skills can be measured using scoring system
25	1999	O'Toole et al	Cambridge USA	20	Novice & Experts	Virtual Reality Error Scoring & Time Open	Difference between experts & novices, improvement with training	VR may be relevant in training & technical skill assessment

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
24	1999	Macmillan et al	Dundee UK	10	Intermediate	Bench Model Motion Analysis, Time & Error Scoring MIS	Good inter-rater reliability, concurrent validity	Good correlation between performance & assessment
23	1999	Qayumi et al	Vancouver Canada	87	Novice	Bench Models Global Open	Improvement in performance of surgical skills	Course improves surgical skills
22	1999	Goff ¹ et al	Seattle USA	24	Intermediate	Bench Models Global & Time Open MIS	Improvement in time & technical skill	Training improved surgical skills
21	1999	Gallagher ¹ et al	Belfast UK	16	Novice	Virtual Reality Error Scoring, Time & Motion Analysis MIS	Training improved performance	VR training may help novices to acquire technical skills
20	1999	Summers et al	Milwaukee USA	69	Novices	Bench Models Global Open	Computer group improved their technical skills	Training can improve skills in novices
19	1999	Derossis et al	Montreal Canada	10	Novice & Intermediate	Bench Models Global & Time MIS	Correlation of grade of surgeon to performance scores	Performance improved with training and grade

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
18	1999	Anastakis ² et al	Toronto Canada	23	Novice	Bench Models Global & Check Lists Open	Training on bench models improves performance compared to text reading	Training improves performance
17	1999	Cauraugh et al	Gainsville USA	6	Novice	Bench Model Global & Time Open	Feedback improved the surgical performance	Mindfulness learning is more effective
16	1998	Taffinder ¹ et al	London UK	16	Novices & Experts	Virtual Reality Motion Analysis, Time & Error Scoring MIS	Construct validity, training improved efficiency and reduction of errors	VR can assess laparoscopic skills & training improves these skills
15	1998	Regeher et al	Toronto Canada	53	Intermediate	Bench Models Global & Check List Open	Reliability good for check list and global, but global was better	Global scoring showed better reliability, construct & concurrent validity
14	1998	Starkes et al	Hamilton Canada	13	Novices	Bench Model Time Open	Improvement in performance	Test able to assess improvement in performance
13	1998	Rogers ¹ et al	Augusta USA	82	Novices	Bench Model Time & Check List Open	No difference in the two types of training	Feedback improved surgical performance

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
12	1998	Joice ¹ et al	Dundee UK	9	Intermediate & Expert	Bench Model Check List, Time & Error Scoring MIS	Performance was dependent on surgeon's style employed	Task accuracy was surgeon dependent
11	1997	Hanna ² et al	Dundee UK	10	Intermediate & Expert	Bench Model Mechanical & Time MIS	Difference in scores for type of knot performed	Scoring system provides reliable assessment of knot security
10	1997	Martin et al	Toronto Canada	20	Novice & Expert	Bench Models Real Animals Global & Check List Open	Bench models & animal operating had good correlation. Global score had good reliability & construct validity. Check list only reliability	Global scoring system has good inter-rater reliability & construct validity
9	1997	Rosser et al	New Haven USA	150	Intermediate & Expert	Bench Models Time & Error Score MIS	Improvement in skills with training	Drills helped in acquiring laparoscopic skills

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
8	1997	Reznick ¹ et al	Toronto Canada	48	Novice & Intermediate	Bench Models Real Animal Global & Check List Open	Inter-rater reliability good & construct validity for global & check list.	Global score has good reliability & construct validity
7	1996	Faulkner et al	Toronto Canada	12	Novice & Intermediate	Bench Models Real Animal Global & Check List Open	Inter-rater reliability better for intermediate than novice	Global score may be valid for intermediates
6	1996	Chak et al	Cleveland USA	19	Novice, intermediate & Expert	Real Patients Time & Check List MIS	Trainees were slower & had poorer performance than experts	Trainees did not achieve expert level after 100 cases
5	1996	Watkins et al	Chicago USA	21	Intermediate	Real Patients Check List MIS	Learning curve varied depending on complexity of task	Trainees achieve competency after 100 cases

Table 2.4 – Papers on Technical Skills Assessment 2005 to 1975

Paper	Year	Authors	Origin	Number of subjects	Type of Subjects	Methods & Materials	Results	Conclusions
4	1994	Winckel et al	Toronto Canada	12	Novice & Expert	Real human Global & Check List Open	Good inter-rater reliability & construct validity for global & check list	Tool could be used as to assess surgical trainees
3	1993	Cass et al	Minneapolis USA	12	Novice & Intermediate	Real human Time & Check List MIS	Technical goals not reached by 100 cases	More than 100 cases are required to achieve competency
2	1992	Lossing et al	Toronto Canada	28	Novices	Real animal Check List Open	Technical skills improved	Structured programme improves technical skills in trainees
1	1989	Risucci ² et al	New York USA	93	Novice, Intermediate & Expert	Real Human Global Open	Poor correlation between self and senior ratings	Cognitive and non-cognitive skills difference may be due to training

Appendix Chapter 3 - Human Error & Reliability

Variable & Constant Errors

The accuracy of error prediction depends very largely on the extent to which the factors resulting in errors are understood. The three major elements in error production are:

- 1/ Nature of the task and its environmental circumstances
- 2/ Mechanisms controlling performance
- 3/ Nature of the individual

Intentions, Actions & Consequences

The notion of intention comprises of two key elements:

- 1/ Expression of the end-state to be attained
- 2/ Indication of the means by which it is to be achieved.

Prior Intention & Intentional Action

All intentional actions have intentions in action but not all intentional actions have prior intentions. Actions without prior intentions fall into two broad classes: intentional and non-intentional actions.

Intentional Actions without Prior Intention

These are spontaneous actions e.g. going to the shops, we do not consciously think of all the small component operations in advance i.e. taking your wallet, putting your coat on, closing and locking the door, walking or driving to the shop etc.

Non-intentional or Involuntary Actions

The term error can only be applied to intentional actions. Error types depend upon two kinds of failure:

1/ Failure of actions to go as intended (slips and lapses)

2/ Failure of intended actions to achieve their desired consequences (mistakes).

Unintended Actions

Actions that deviate from intention fall in to two classes: those that nevertheless achieve their intended goal and those that do not. Absent-mindedness occurs in humans when we become aware that our actions have strayed from their intended path. These occur as a result of the performance of a largely automated task in a familiar environment and a marked degree of attentional 'capture' by something other than the task in hand (Reason 1979, Norman 1981).

Intended Actions & Mistakes

Even when the intended actions proceed as planned, they can still be judged as erroneous if they fail to achieve their intended outcome. The two basic error forms are due to either a planning error (mistakes) or execution failures (slips and lapses).

Table 3.5 – Surgical Task & Error Analysis Papers 1995 - 2005

Author	Methods	Setting	Subjects	Number of Cases	Results/Conclusions
Cao ¹ et al 1996	Task analysis of specific tasks & motion analysis of generic movements	Workshop Live animals	5 Novice 1 Expert	17-38	Novices performed specific tasks and generic movements slower than expert surgeon
Joice ² et al 1998	Task & error analysis of specific & generic aspects of laparoscopic cholecystectomy	Theatre Live human	8 Intermediate	20	Observational data capture can study human error in laparoscopic surgery. Categorization of interstep (procedural) & intrastep (execution) errors
Cao ² et al 1999	Hierarchical task analysis of laparoscopic cholecystectomy, inguinal hernia repair, fundoplication	Workshop & Theatre Live animal & human	1 Intermediate 4 Expert	Not Known	HTA can be used to study the relationship between goals & actions at various levels of the hierarchy
Eubanks et al 1999	Error checklist scoring of procedural tasks in laparoscopic cholecystectomy	Theatre Live human	28 Intermediate 2 Expert	30	Error checklist scoring was reliable, valid & practical
Mehta et al 2002	Generic sub-task analysis of instruments used in various laparoscopic operations	Theatre Live human	Not Known	29	Laparoscopic instruments used to perform a variety of maneuvers in addition to primary function
Tang ¹ et al 2004	Consequential & inconsequential errors derived from task analysis of laparoscopic cholecystectomy	Theatre Live human	26 Intermediate	200	OCHRA provides objective & comprehensive assessment of surgical performance
Seymour ² et al 2004	Error event definitions of generic & specific aspects of laparoscopic cholecystectomy	Theatre Live human	14 Intermediate	14	Reliability of defined procedural errors

Appendix Chapter 4 - Hierarchical Task Analysis in Surgery

Table 4.13 - Task Analysis of Inguinal Hernia – Anaesthetist `s View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
1	Prepare patient in anaesthetic room	Do subtasks 1.1 to 1.5	1.1 1.2 1.3 1.4 1.5	Check consent Check patient details Attach monitors Gain IV access Pre-oxygenate if required	Correct paperwork if incorrect Reassess patient if abnormal monitor reading	Performing subtasks 1.1 to 1.5 Assisting anaesthetist
2	Induction of anaesthesia	Do subtasks 2.1 to 2.2	2.1 2.2	I.V. (in most patients) or inhalational if indicated Other drugs as necessary	Adverse drug reaction treated as necessary	Assisting anaesthetist
3	Airway control	Do subtasks 3.1 to 3.3	3.1 3.2 3.3	Maintain oxygenation and mask ventilate if necessary Insert airway or intubate as required Secure airway and check its position	Correcting airway obstruction and oxygenating via facemask if unable to intubate. Call for assistance and equipment if difficult intubation	
4	Further anaesthetic intervention as required	Do subtasks 4.1 to 4.3	4.1 4.2 4.3	Large bore IV if not placed at 1.4 Eye care Local nerve block if desired		
5	Transfer and positioning of patient	Do subtasks 5.1 to 5.5	5.1 5.2 5.3 5.4 5.5	Maintain clinical observation of patient during transfer Safe transfer to operating table using lifting device if required Reinstate monitoring Check pressure areas Supine position for surgery	314 314	ping with sfer and monitoring of patient

Table 4.13 - Task Analysis of Inguinal Hernia – Anaesthetist `s View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
6	Lead up to start of surgery	Do subtasks 6.1 to 6.3	6.1 6.2 6.3	Check monitoring and vital signs Check IV access site Check adequacy of anaesthesia	If abnormal vital signs check equipment and patient. Resite IV if necessary Deepen anaesthesia if patient 'light'	
7	During surgery	Do sub tasks 7.1-7.3	7.1 7.2 7.3	Continuous monitoring Record keeping Looking out for vagal response to pulling on spermatic cord	Give atropine if necessary	
8	Post surgery	Do sub tasks 8.1-8.2	8.1 8.2 8.2	Reverse relaxant if they have been used Turn off anaesthetic delivery when op finished Ensure adequate analgesia has been given before patient wakes	Give analgesia as required	
9	Emergence	Do sub tasks 9.1-9.4	9.1 9.2 9.3 9.4	Continue full monitoring Remove airway when patient awake and protecting own airway Supply supplemental O2 Transfer to appropriate trolley and take to recovery unit	reintubate if necessary if severe laryngo-spasm	Help with safe transfer
10	Recovery	Do sub tasks 10.1-10.4	10.1 10.2 10.3 10.4	Continue full monitoring Verbal handover to nurse Ensure satisfactory cardiorespiratory parameters Ensure adequate analgesia	Adjust airway and fluid management to correct	
11	Administration		11.1	Ensure notes completed and prescriptions		

Table 4.14 - Task Analysis of Inguinal Hernia – Scrub Nurse’s View

No.	Task	Plan	No.	Sub-tasks	Recovery	What runner should be doing
1	Instrument & Equipment Preparation	Do subtasks 1.1 to 1.5	1.1 1.2 1.3 1.4 1.5	Consent Correct Positioning Appropriate diathermy attachment & check Pre check Instruments & swabs Ensure Availability & equipment is in full working order	Correct or inform any irregularities appropriately If mesh, materials, instrument and equipment not available request them again and replace faulty items	See subtask 1.1 to 1.5 Safe transfer on to table
2	Assisting surgeon during operation	Do subtasks 2.1 to 2.3 continuous and repetitive during procedure	2.1 2.2 2.3	Drape patient Instruments and materials try to anticipate at all times Maintain safety	Drape dropped or becomes non-sterile, replace Equipment or material dropped, ask for replacement Torn surgical glove, replace etc	See subtask 2.1-2.3 Anticipates surgeon and scrub nurses needs
3	Operation nearing the end	Do subtasks 3.1 to 3.4	3.1 3.2 3.3 3.4	First instrument and swab count check Post closure checking swabs & instruments Local anaesthetic if required Dressings	If incorrect recheck If incorrect recheck, inform surgeon If incorrect or dropped, replace Replace if incorrect or dropped	See subtask 3.1 -3.3
4	Transfer to bed safely	Do subtasks 4.1 to 4.3	4.1 4.2 4.3	Patient’s operated area clean Diathermy pad removed and check site Transfer patient	If dirty, clean Inform if diathermy burn	Transfer to bed safely
5	Transfer to recovery & handover to recovery	Do subtasks 5.1 to 5.3	5.1 5.2 5.3	Notes available Forms done Transfer & handover	If incorrect, correct and complete	Transfer to recovery Safely

Table 4.15 - Task Analysis of Long Saphenous Varicose Veins – Anaesthetist`s View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
1	Prepare patient in anaesthetic room	Do subtasks 1.1 to 1.5	1.1 1.2 1.3 1.4 1.5	Check consent Check patient details Attach monitors Gain IV access Pre-oxygenate if required	Correct paperwork if incorrect Reassess patient if abnormal monitor reading	Performing subtasks 1.1 to 1.5 Assisting anaesthetist
2	Induction of anaesthesia	Do subtasks 2.1 to 2.2	2.1 2.2	I.V. (in most patients) or inhalational if indicated Other drugs as necessary	Adverse drug reaction treated as necessary	Assisting anaesthetist
3	Airway control	Do subtasks 3.1 to 3.3	3.1 3.2 3.3	Maintain oxygenation and mask ventilate if necessary Intubate with endotracheal tube if prone position anticipated. LMA if supine Secure airway and check its position	Correcting airway obstruction and oxygenating via facemask if unable to intubate. Call for assistance and equipment if difficult intubation	
4	Further anaesthesia intervention as required	Do subtasks 4.1 to 4.3	4.1 4.2 4.3	Large bore IV if not placed at 1.4 Eye care Additional eye pads if prone position		

Table 4.15 - Task Analysis of Long Saphenous Varicose Veins – Anaesthetist 's View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
5	Transfer and positioning of patient	Do subtasks 5.1 to 5.6	5.1	Maintain clinical observation of patient during transfer		Helping with transfer and monitoring of patient
			5.2	Safe transfer to operating table using lifting device if required		
			5.3	Reinstate monitoring		
			5.4	Check pressure areas		
			5.5	Check no traction on brachial plexus if prone.		
			5.6	Warming devices applied		
6	Lead up to start of surgery	Do subtasks 6.1 to 6.3	6.1	Check monitoring and vital signs	If abnormal vital signs check equipment and patient. Resite IV if necessary Deepen anaesthesia if patient 'light'	
			6.2	Check IV access site		
			6.3	Check adequacy of anaesthesia		
7	During surgery	Do sub tasks 7.1-7.5	7.1	Continuous monitoring	Increase IV fluids as necessary continuing management of hypovolaemia if required Treat cardio vascular and respiratory response to this Deepen anaesthesia and/or give analgesia	
			7.2	Record keeping		
			7.3	Looking out for bleeding		
			7.4	Head down tilt		
			7.5	Look out for increased stimulation during stripping of vein		

Table 4.15 - Task Analysis of Long Saphenous Varicose Veins – Anaesthetist 's View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
8	Post surgery	Do sub tasks 8.1-8.3	8.1 8.2 8.3	Reverse relaxant if they have been used Turn off anaesthetic delivery when op finished Ensure adequate analgesia has been given before patient wakes	Give analgesia as required	
9	Emergence	Do sub tasks 9.1-9.4	9.1 9.2 9.3 9.4	Continue full monitoring Remove airway when patient awake and protecting own airway Supply supplemental O2 Transfer to appropriate trolley and take to recovery unit	Reintubate if necessary if severe laryngo-spasm	Help with safe transfer
10	Recovery	Do sub tasks 10.1-10.4	10.1 10.2 10.3 10.4	Continue full monitoring Verbal handover to nurse Ensure satisfactory cardio-respiratory parameters Ensure adequate analgesia	Adjust airway and fluid management to correct Give IV/IM opiates PRN	
11	Administration		11.1	Ensure notes completed and necessary prescriptions made		

Table 4.16 - Task Analysis of Sapheno-Femoral Junction Ligation – Scrub Nurse

No.	Task	Plan	No.	Sub-tasks	Recovery	What runner should be doing
1	Instrument & Equipment Preparation	Do subtasks 1.1 to 1.5	1.1 1.2 1.3 1.4 1.5	Consent Correct Positioning Appropriate diathermy attachment & check Pre check Instruments & swabs Ensure Availability & equipment is in full working order	Correct or inform any irregularities appropriately If vascular instruments, materials, routine instrument and equipment not available request them again and replace faulty items	See subtask 1.1 to 1.5 Safe transfer on to table
2	Assisting surgeon during operation	Do subtasks 2.1 to 2.3 continuous and repetitive during procedure	2.1 2.2 2.3	Drape patient Instruments and materials try to anticipate at all times Maintain safety	Drape dropped or becomes non-sterile, replace Equipment or material dropped, ask for replacement Torn surgical glove, replace Bleeding, set up suction	See subtask 2.1-2.3 Anticipates surgeon and scrub nurses needs
3	Operation nearing the end	Do subtasks 3.1 to 3.4	3.1 3.2 3.3 3.4	First instrument and swab count check Post closure checking swabs & instruments Local anaesthetic if required Dressings	If incorrect recheck If incorrect recheck, inform surgeon If incorrect or dropped, replace Replace if incorrect or dropped	See subtask 3.1 -3.3

Table 4.17 - Task Analysis of Laparoscopic Cholecystectomy – Anaesthetist `s View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
1	Prepare patient in anaesthetic room	Do subtasks 1.1 to 1.5	1.1 1.2 1.3 1.4 1.5	Check consent Check patient details Attach monitors Gain IV access Pre-oxygenate if required	Correct paperwork if incorrect Reassess patient if abnormal monitor reading	Performing subtasks 1.1 to 1.5 Assisting anaesthetist
2	Induction of anaesthesia	Do subtasks 2.1 to 2.2	2.1 2.2	I.V. (in most patients) or inhalational if indicated Other drugs as necessary	Adverse drug reaction treated as necessary	Assisting anaesthetist
3	Airway control	Do subtasks 3.1 to 3.4	3.1 3.2 3.3 3.4	Maintain oxygenation and mask ventilate if necessary Intubate with endotracheal tube Secure airway and check its position Ensure adequate ventilation and oxygenation	Correcting airway obstruction and oxygenating via facemask if unable to intubate. Call for assistance and equipment if difficult intubation	
4	Further anaesthetic intervention as required	Do subtasks 4.1 to 4.3	4.1 4.2 4.3	Large bore IV if not placed at 1.4 Eye care Place NG tube		

Table 4.17 - Task Analysis of Laparoscopic Cholecystectomy – Anaesthetist `s View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
5	Transfer and positioning of patient	Do subtasks 5.1 to 5.6	5.1	Maintain clinical observation of patient during transfer		Helping with transfer and monitoring of patient
			5.2	Safe transfer to operating table using lifting device if required		
			5.3	Reinstate monitoring		
			5.4	Check pressure areas		
			5.5	Supine position for surgery		
			5.6	Warming devices applied		
6	Lead up to start of surgery	Do subtasks 6.1 to 6.3	6.1	Check monitoring and vital signs	If abnormal vital signs check equipment and patient.	
			6.2	Check IV access site	Resite IV if necessary	
			6.3	Check adequacy of anaesthesia	Deepen anaesthesia if patient 'light'	
7	During surgery	Do sub tasks 7.1-7.7	7.1	Continuous monitoring	Ensure stomach de-compressed.	
			7.2	Record keeping	Give atropine if vagal response	
			7.3	Looking out for response to pneumo-peritoneum	Look out for CO2 embolus and treat as necessary	
			7.4	Head up tilt	Treat cardio-vascular response to this	
			7.5	Ensure adequate muscle relaxation	Monitor with NMJ monitor	
			7.6	Give IV antibiotics		
			7.7	Stop ventilating if cholangiogram		

Table 4.17 - Task Analysis of Laparoscopic Cholecystectomy – Anaesthetist 's View

No.	Task	Plan	No.	Sub-tasks	Recovery	What ODP should be doing
8	Post surgery	Do sub tasks 8.1-8.3	8.1 8.2 8.2 8.3	Reverse relaxant if they have been used Turn off anaesthetic delivery when op finished Ensure adequate analgesia has been given before patient wakes Remove NG tube	Give analgesia as required	
9	Emergence	Do sub tasks 9.1-9.4	9.1 9.2 9.3 9.4	Continue full monitoring Remove airway when patient awake and protecting own airway Supply supplemental O2 Transfer to appropriate trolley and take to recovery unit	reintubate if necessary If severe laryngo-spasm	Help with safe transfer
10	Recovery	Do sub tasks 10.1-10.5	10.1 10.2 10.3 10.4 10.5	Continue full monitoring Verbal handover to nurse Ensure satisfactory cardio-respiratory parameters Ensure adequate analgesia Give or prescribe anti-emetics	Adjust airway and fluid management to correct Give IV/IM morphine PRN	
11	Administration		11.1	Ensure notes completed and necessary prescriptions made		

Table 4.18 - Task Analysis of Laparoscopic Cholecystectomy – Scrub Nurse’s View

No.	Task	Plan	No.	Sub-tasks	Recovery	What runner should be doing
1	Instrument & Equipment Preparation	Do subtasks 1.1 to 1.5	1.1 1.2 1.3 1.4 1.5	Consent Correct Positioning Appropriate diathermy attachment & check Pre check laparoscopic instruments & swabs, CO2 and camera Ensure Availability & equipment is in working order	Correct or inform any irregularities appropriately If materials, instrument and equipment not available request them again and replace faulty items	See subtask 1.1 to 1.5 Safe transfer on to table
2	Assisting surgeon during operation	Do subtasks 2.1 to 2.3 continuous and repetitive during procedure	2.1 2.2 2.3	Drape patient Instruments and materials try to anticipate at all times Maintain safety	Drape dropped or becomes non-sterile, replace Equipment or material dropped, replace Torn surgical glove, replace Bleeding, set up suction irrigator	See subtask 2.1-2.3 Anticipates surgeon and scrub nurses needs
3	Operation nearing the end	Do subtasks 3.1 to 3.4	3.1 3.2 3.3 3.4	First instrument and swab count check Post closure checking swabs & instruments Local anaesthetic Dressings	If incorrect recheck If incorrect recheck, inform surgeon If incorrect or dropped, replace Replace if incorrect or dropped	See subtask 3.1 -3.3
4	Transfer to bed safely	Do subtasks 4.1 to 4.3	4.1 4.2 4.3	Patient’s operated area clean Diathermy pad removed and check site Transfer patient	If dirty, clean Inform if diathermy burn	Transfer to bed safely
5	Transfer to recovery & handover to recovery nurse	Do subtasks 5.1 to 5.3	5.1 5.2 5.3	Notes available Forms completed Transfer and handover to recovery	If incorrect, correct and complete	Transfer to recovery Safely

Appendix Chapter 7

Assessment of Technical Skills in Open Live Inguinal Hernia Surgery

Appendix 7.1

There are two intra-class correlation coefficients, a single and an average measure. Single is an index of reliability of the scores for a typical, single observer. One employs it when you are going to collect most of your data using only one observer at a time. When a researcher decides that there is too much uncertainty to use an individual rating and/or when the unit of analysis is the average of both ratings, the average should be used. Therefore we use this for reliability of the global scores.

Appendix 7.2

The intra-class correlation coefficients are applied when variables are considered continuous. When summing the scores to compute an overall coefficient in the Global Technical Skills Assessment (even if the scale is categorical), the resulting values have around 15 orderings, which is acceptable for these variables to be considered continual. In the case of the Checklist, because of the original variables are of the type Yes/No or Ok/Yes(corrected)/Yes(uncorrected), there are too few categories to be considered continuous. Therefore, to assess reliability we use Weighted Kappa coefficients.

Appendix 7.3

Weighted Kappa Coefficient

When there are more than two categories and they are ordered, a difference between raters of, let's say, two categories would indicate more disagreement than if they disagree on just one category. This feature is incorporated in the Weighted Kappa Coefficient.

A general interpretation for the Kappa coefficient scores are as follows:

Kappa Coefficient	Reliability
< 0.00	Poor
0.00 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 0.99	Almost Perfect
1.00	Perfect