A
Virtual University Infrastructure
For
Orthopaedic Surgical Training
With
Integrated Simulation

SUBMITTED BY

MR SIMON ANDRĒ WELHAM GRANGE
MB ChB (BRISTOL) FRCS (Tr & Orth) (Eng)

TO THE UNIVERSITY OF EXETER AS A THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN ENGINEERING AND COMPUTER SCIENCE, JUNE 2005

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or another university.

……………………………………….(Signature)

ABSTRACT

This thesis pivots around the fulcrum of surgical, educational and technological factors. Whilst there is no single conclusion drawn, it is a multidisciplinary thesis exploring the juxtaposition of different academic domains that have a significant influence upon each other. The relationship centres on the engineering and computer science factors in learning technologies for surgery. Following a brief introduction to previous efforts developing surgical simulation, this thesis considers education and learning in orthopaedics, the design and building of a simulator for shoulder surgery. The thesis considers the assessment of such tools and embedding into a virtual learning environment. It explains how the performed experiments clarified issues and their actual significance. This leads to discussion of the work and conclusions are drawn regarding the progress of integration of distributed simulation within the healthcare environment, suggesting how future work can proceed.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACCOMPANYING MATERIAL

ICT Skills for Healthcare Professionals

Technical Report Number: ECSTR-IAM02-004 dated 08 July 2002
IAM Research Group
Department of Electronics and Computer Science
University of Southampton
SOUTHAMPTON
SO17 1BJ

British Library Cataloguing in Publication Data
ISBN: 0-854-32773-8

AUTHOR'S DECLARATION - ACKNOWLEDGEMENTS

This work is submitted as a PhD thesis for the department of engineering and computer science in Exeter, United Kingdom (UNIV OF EXETER) and has been partly funded by the European Commission (EC) to develop a web site, *VOEU©*, for the study of Orthopaedics for both basic and also higher surgical trainees and trained practising surgeons. The influence of Prof. Philippe Cinquin (Grenoble) and Dr Klaus Radermacher (Aachen) are distinctive. My supervisors, Drs. Steve Turner, Gareth Jones and Prof. Ajit Narayanan can be credited with changing the way I think. Though this is my own work spread over several years, eclectically adopting the ideas of many different domains, it would not have been possible without the help and support of the department of Electronics and Computer Science in Southampton University (SOTON), in particular, the assistance of Prof. Wendy Hall, Prof. David De Roure, Dr Gary Wills and Dr Les Carr. Much of the educational input has come from the VOEU project and the contributions of Prof. Grainne Conole from the department of education. Finally I wish to acknowledge the discussion with the Dr Oliver Tonetti at the Scuola Superiore Sant' Anna, Pisa regarding the evaluation of their simulation system, using the scoring system that I proposed, currently submitted for publication. There is of course my family, Julie and Sophie, who have had to compromise in many ways to accommodate this thesis, and the senior orthopaedic surgeons in the Wessex region, who have tolerated this apparent diversion from the conventional orthopaedic apprenticeship. In particular I have had considerable guiding support from Mr Tim Bunker at the Princess Elizabeth Orthopaedic Centre in Exeter.

ACRONYMS

AH	Adaptive Hypermedia
AIPES	Autonomy, Interactivity, Presence, Environment & Scenario
ATM	Asynchronous Transfer Mode
BST	Basic Surgical Trainee
CAMIS	Computer Assisted Minimally Invasive Surgery
CAS	Computer Aided Surgery
CBL	Computer-Based (distance) Learning
CCST	Certificate of Completion of Specialist Training
CCT	Certificate of Completion of Training
CME	Continuing Medical Education
CPD	Continuing Professional Development
CRT	Cathode Ray Tube
CT	Computed Tomography (medical imaging)
DARPA	Defence Advanced Research and Procurement Agency
DOH	Department of Health
DRJ	Dynamic Review Journal
DTD	Document Type Definition
ECDL	European Computer Driving Test
EM	Electromagnetic
EPSRC	Engineering & Physical Sciences Research Council
EVW	Exeter Virtual Worlds (SIM1 + VOU1)
EWTD	European Working Time Directives
FOV	Field Of View
FZK	Forschungszentrum Karlsruhe
GMC	General Medical Council
GUI	Graphical User Interface
GUIDE	Graphical User Interface Design & Evaluation
HCI	Human Computer Interface
HDTV	High Definition TeleVision
HIA	Helmholz Institute Aachen (Germany)

HIS	Hospital Information Systems
HIV	Human Immunodeficiency Virus
HPL	Human Performance & Limitations
HST	Higher Surgical Trainee
HTML	HyperText Mark-up Language
IAM	Intelligence, Agents, Multimedia (Research Group)
ICT	Information & Communications Technology
ID	Instructional Design
IGOS	Image Guided Orthopaedic Surgery
ILA	Individual Learning Agreement
IMS	Information Management System
IP	Internet Protocol
JAA	Joint Aviation Authority
JCHST	Joint Committee for Higher Surgical Training
JPEG	Joint Photographic Experts Group (Still image format)
MAS	Minimal Access Surgery
MEOM	Multimedia Educational Orthopaedic Module
MIS	Minimally Invasive Surgery
MIST	Minimally Invasive Surgical Trainer
MPEG	Moving Photographic Experts Group (Video image format)
MRI	Magnetic Resonance Imaging
NLP	Neurolinguistic Programming
NHS	National Health Service (United Kingdom)
PMETB	Postgraduate Medical Education Training Board
OLE	Object linking and embedding (Microsoft™)
OSATS	Objective Structured Assessment of Technical Skills
OSML	Orthopaedic Surgical Mark-up Language
RAM	Random Access Memory
RCS	The Royal College of Surgeons of England
RITA	Regional In-service Training Agreement
SSSA	Scuolo Superiore Sant' Anna (Italy)
SIM1	Simulation 1 (Exeter Virtual Worlds)
SIM2	Simulation 2 (University of Bristol)

SIM3	Simulation 3 (University of Southampton)
SGML	Structured General Mark-up Language
SOA	Service Oriented Architecture
TER	Transfer Effectiveness Ratio
TQM	Total Quality Management
UI	User Interface
URL	Uniform Resource Locator
VE	Virtual Environment
VHP	Visible Human Project
VOEU	Virtual Orthopaedic European University
VOU1	Virtual Orthopaedic University 1 (University of Exeter)
VOU2	Virtual Orthopaedic University 2 (University of Southampton – VOEU)
VR	Virtual Reality
V&V	Verification and Validation
WIMP	Windows, Icons, Menus, Pointers
WUN	World Universities Network

1 Introduction

1.1 Introduction

Figure 1.1 Sir William Osler (1849-1919)

William Osler (Figure 1.1) wrote in ‘The Principles and Practice of Medicine’ (1) in 1892 that:

‘To learn medicine without books is to sail an uncharted sea,

While to learn medicine only from books, is to not go to sea at all’.

Progressing from the 19th century to the 21st century, with reference *en passant* to the 20th, this thesis demonstrates how the changes in technologies are apparent, though the necessary change in philosophy to employ them remains tardy. The thesis retains the spirit of Osler, whilst illustrating how novel techniques in computer science can be applied to augment the educational process through the enhancement of multimedia and hypermedia in surgical training. This chapter lays down the objective of an integrated simulation and education system, and sets the cultural context for the thesis. The timely provision of surgical training is significant due to the reduction in hours available now to train surgeons and aspiration for consistent standards, especially when the UK and Spain are training non-medical staff to operate. The thesis goes on to discuss the core reasons for developing surgical simulation, outlining the political demands leading to such simulation. These have resulted in the adaptation of the current computing technologies to partially fill the niche in medical education.

1.2 Thesis Statement

The changes in social, political, technical, and financial factors led to a perceived crisis of confidence in the standards of medical competency in United Kingdom Healthcare professionals in the latter half of the 1990's (2). This led to the potential for development of new education systems that represent both greater flexibility and also greater accountability[1]. With the introduction of the 'European Working Time Directive' legislation, there is a potential reduction of the hours available for training to 6,000 from 25,000 at present, already reduced from the 45,000[2] of the traditional apprenticeship. Novel approaches to enhance the process of surgical training, including infrastructure, will become essential.

To focus the thesis upon answering the key question of the suitability of developing such an infrastructure the following thesis statement was written. *Surgical training benefits from the integration of procedural trainers into a declarative learning environment managed using a distributed architecture.*

1.3 Hypothesis

Leading from the thesis statement, the hypothesis to test is: '*A virtual university infrastructure offers the potential of providing an effective distributed learning environment for orthopaedic surgeons'*. Simulation is but one medium for the conveyance of information to learners. The author proposes that the introduction of surgical simulation using video-based systems will improve the learning of shoulder arthroscopy if incorporated within an integrated training module that employs a User Interface (UI) to control an input device for the diagnostic arthroscopy trainer.

An iterative design philosophy has been developed. Initiated in the University of Exeter, then continued through the University of Bristol, to the University of Southampton, the simulations have been standardised to SIM1-3 respectively. The Virtual

---

[1] The 'Bristol' Case http://www.gmc-uk.org/news/archive/bristol.htm relating to failure to provide adequate safeguards with respect to the provision of paediatric cardiac surgery and the 'Harold Shipman' case http://www.doh.gov.uk/hshipmanpractice/ where a General practitioner was found guilty of murdering his patients both highlighted the loss of public confidence in the medical profession during the 1990's in the United kingdom and raised questions regarding competency.

[2] Discussed at The Royal Society of Medicine 2002 meeting upon the future of surgical training summary of current practice. http://www.rsm.ac.uk

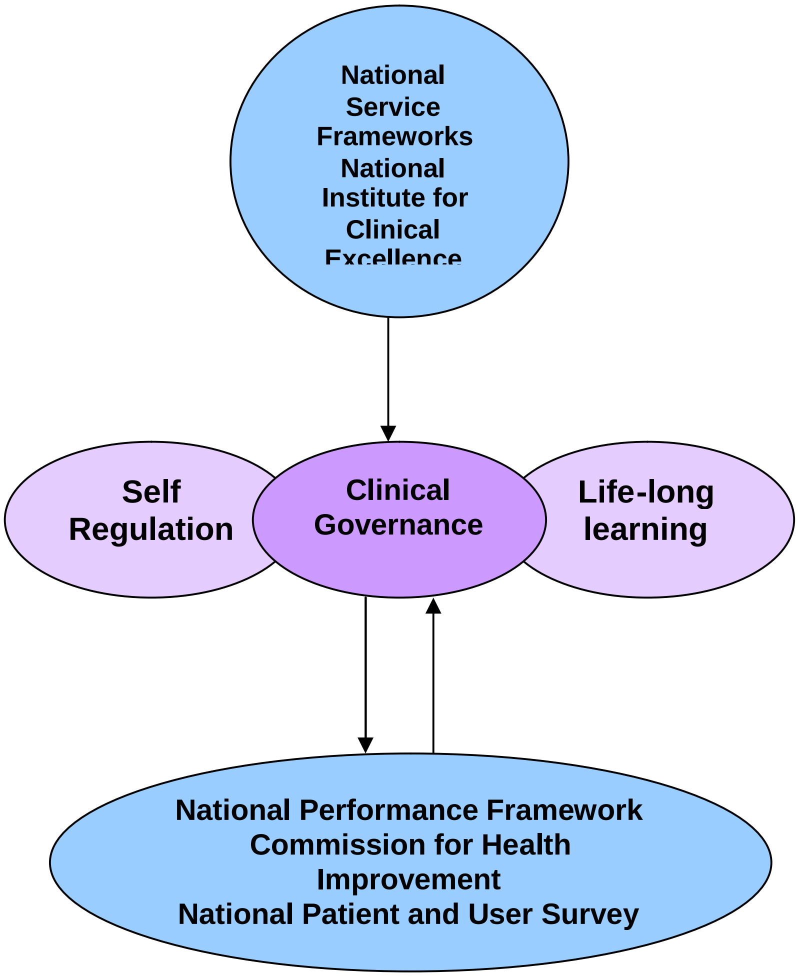
Orthopaedic University Infrastructures (VOUs) have been termed VOU1 and VOU2. The development of the Exeter Virtual Worlds (EVW) simulation system and also the Exeter Virtual Orthopaedic University (The 1st design phase simulation and the subsequent model) and subsequently the Virtual Orthopaedic European University (VOU2) and the 3rd generation simulation (SIM3) used in trials must be seen within its chronological, cultural, and technological contexts for the thesis which spans 1997-2004. Evaluation of the quality of experience is quantifiable, by means of:

- Usability testing of the user interface
- User perceived educational value of simulation.

The area of surgery chosen to explore this hypothesis is that of shoulder arthroscopy using techniques explained in detail below. The thesis statement argues that the training of shoulder arthroscopy within structured specialist surgical courses is improved by the novel introduction of computer-based surgical simulation, which can be managed effectively using a client server architecture that offers a significant advantage over the resources that were previously available.

1.4 Political significance

The computer applications described in this thesis are designed to meet specific user requirements. The description of the training framework into which the simulation system will ultimately be integrated is of particular importance for its socio-cultural acceptance and adoption. For this reason, and to assist those readers who are involved in developing future training systems, an overview of the current framework is provided below on surgical ontology development. The current *National Health Service (NHS*) political directive is for the 'development of better healthcare for all' through the model of the 'First Class Service' (3) (Figure 1.2). This has provided an overarching structure for future process management embodying both *Total Quality Management (TQM)* and *Clinical Skills (Quality Management)* under the umbrella of *clinical governance* (4).



CLEAR STANDARDS OF SERVICE

DEPENDABLE LOCAL DELIVERY

MONITORED STANDARDS

Figure 1.2 - The management of quality in the NHS - from "A First Class Service'

In order to assure dependable local delivery meeting national standards the aim is to ensure standards are monitored through both internal 'peer' review (professional self-regulation) and external review such as community health initiatives.

1.5 Core reasons for developing surgical simulation

The quality management system will assist trainees in the auditing of results, as already proposed by the preparation of clinical logbooks(5). The value of simulation in training will only become truly apparent once established and accepted, though it is possible currently to demonstrate how simulators can benefit trainees through trials such as those described in Chapters 4 and 8. Whilst the aims and objectives of an integrated simulation and education system will vary on a regional basis they will need to fulfil national recommendations including mapping onto the syllabus and curriculum. The cultural context will have a significant impact upon its acceptance. A summary of the thesis chapters is presented below.

1.6 Chapter 2 - Education and learning in Orthopaedics

By covering prior research into surgical simulation, this chapter outlines the historical perspective that has preceded the current research into the development of computer-based surgical simulation. It then explores prior research on surgical simulation, using computer simulated *Virtual Environments (VEs)* in more detail. Defining the *Instructional Design (ID)* is the basis of developing a new educational strategy. Relevant to surgeons and computer scientists is the assertion that the pedagogy must encompass the developing multimedia integration strategy to accommodate the syllabus for specific populations of users preparing for specific examinations, such as *the Certificate Completion of Specialist Training* (CCST) examinations. The system user requirements stem from this.

1.7 Chapter 3 - Shoulder simulator design and development

The basic physiological targets that constitute human performance limitations are set, and thus the thresholds that VE needs to achieve in order to appear *real* to the user are defined. This chapter qualifies the appropriate thresholds that impact upon system design. Considering the principles of *Human Performance and its Limitations (HPL),* the lessons learned from the vehicle-operator training industry are applied, suggesting a suitable framework for proto-standards in surgery. The stages of evolution of prototypes are detailed.

1.8 Chapter 4 - The evaluation of a shoulder simulator

To assess the approaches adopted in surgical simulation, the simulation environment needs to be evaluated formally using a scoring system that considers more than just the virtual environment itself. The Autonomy, Interactivity, Presence, Environment & Scenario (AIPES) scoring system reflecting the major aspects of simulator performance was born. This was devised specifically for the purpose of surgical simulation evaluation, attempting also to encompass other simulation techniques and methods. It relates the fundamental features of a simulation (autonomy, interactivity and presence), with the context dependent features that are particularly relevant to virtual university distributed architectures. These include the learning environment and the educational scenario. The results of an initial usability study, carried out to support the work and comparison with testing on another simulator, provides a foundation for *Validation* and *Verification* in this field. It offers a framework for the evolution of formal evaluation methodologies that can be refined as feedback from users becomes available to close the audit loop, detailing the simulation's actual effectiveness, not just that implied by research evaluation studies.

1.9 Chapter 5 - Surgical and educational factors for a distributed learning environment in orthopaedics

Users will only adopt a virtual learning environment system if the issues of usability are addressed, so the feasibility of employing new media will depend upon its effectiveness in communicating the knowledge that users require. With each design cycle the specifications for the system are set. What can be implemented will of course depend upon the affordability of the new technologies.

1.10 Chapter 6 - The requirements and design of a virtual university infrastructure for learning orthopaedics

The aims of this chapter are to ensure access to all relevant educational material with the development of tailored courses for individuals, offering a customised view of the multimedia knowledge to which they refer.

This chapter links into issues discussed in the previous chapters by ensuring that the surgeon is able to incorporate the learning material into their working practice. They can do this by accessing material only relevant to their present situation through the use of structured learning agreements. These are to be used as the framework for building a surgical ontology.

1.11 Chapter 7 - Virtual university infrastructure facilities

Each section of the Virtual university environment is discussed in detail. This includes library, virtual classroom, and dynamic review journal (DRJ) modules for the presentation of new and established knowledge. It also incorporates simulation and logbook information for the enhancement of these functions. Finally the administration using adaptive hypermedia is discussed.

1.12 Chapter 8 - An evaluation of the virtual university infrastructure as an effective learning environment

This chapter includes analysis of user-centred design and its significance to providing interfaces dedicated to surgical users, by way of reference to standard models for User Interface (UI) design including ‘frames and style sheets’. The interfaces dictate *Staff Training Protocols*, so that setting up new systems will require training of the staff to an adequate standard to operate the UIs.

1.13 Chapter 9 – Conclusion

The final chapter starts with a discussion of the results of introducing a web-based (flexible and distance learning) training environment to surgery. By outlining the proposed methods for the introduction of the web-based training system in conjunction with client-server architecture of the simulator system, it becomes possible to incorporate the new technologies within our major central institutional educational infrastructures. The chapter reviews the current state of development of simulation systems in more detail, emphasising the strengths and weaknesses of alternative simulation systems, which are currently available, compared to the ideal parameters for a surgical simulation medium.

1.14 Summary of chapter 1

This chapter outlined the objective of an integrated simulation and education system, including the cultural context for the thesis. The thesis goes on to discuss the core reasons for developing surgical simulation, outlining the political demands leading to such simulation. This led to the hypothesis: *‘A virtual university infrastructure offers the potential of providing an effective distributed learning environment for orthopaedic surgeons’*, which is defended using arguments supported by experimental results of the usability testing of the user interface and perceived educational value of simulation.

The description of the training system into which the simulation system will ultimately to be integrated is of particular importance for its socio-cultural acceptance and adoption.

2 *Education in Orthopaedics*

2.1 Educating and learning in surgery

The word "*Doctor*" originally meant "*teacher*" derived from the past participle stem *doct-* of the Latin verb *docere* (to teach). The use of *Information & Communication Technologies (ICT)* gives us an opportunity to improve the efficiency of both our teaching and our own learning in the context of Life-long learning also known as *continuous medical education (CME)* and *continuous professional development (CPD).* Basically, they convey the concept of an evolving individual knowledge base and the terms are used interchangeably throughout this thesis.

2.2 Life-long learning in the maintenance of clinical quality

There are three main areas where improvements in healthcare are to be generated:

1. The learning and working environment setting & delivery
2. Maintenance of standards
3. Monitoring of these standards

On an individual level, Life-long learning is a means of maintaining personal standards for all health care workers. To quote the Department of Health:

"Life-long learning will provide NHS staff with the opportunity to continuously update their skills and knowledge to offer the most modern, effective and high quality care to patients." (6)

Currently to accommodate the clinical duties a structure has evolved where surgeons attend either specific courses for 1 to 3 days in duration as an integral part of the CME / CPD infrastructure, or perhaps attend local (Regional) level training meetings for one or two sessions at a time. Trainees aim to successfully attain certification for completion of specialist training (CCST), although this is shortening to become certification of completion of training (CCT) after four years in post instead of 6 years. This is the background to the development of the virtual university support for clinical performance standards and maintenance in line with clinical governance processes (Figure 1.2).

With regard to obtaining ICT skills, it is currently the responsibility of individuals to secure this training for themselves. The options are *'self learning'*, *'in* *house'* or external training courses such as those preparing for the European Computer Driving Test (www.ecdl.co.uk). Life-long learning is the key, as with other aspects of medical education. European standards (7) have been set for healthcare professionals.

2.2.1 Potential benefits of surgical simulation systems

Users can practice and hone their skills repeatedly so as to achieve consistency of performance. Exposure to repetition over a short time frame represents timesaving and allows users to refresh skills prior to actual surgery. The strength is in allowing surgeons of limited experience to gain exposure to tasks that might otherwise be considered too risky in real surgery. This is in fact made easier initially by the over- simplification of the tasks and limited subtask modelling. This is becoming increasingly important as suitable training cases are being relocated to non-training facilities for political reasons.

2.2.2 Potential pitfalls of surgical simulation systems

The weaknesses of such systems may relate to the fact that the early simulators were not strictly part of the instructional design plan and therefore are not clearly mapped to learning outcome objectives within specific syllabi. This impacts upon validation and verification. Criterion-referenced measurements such as anatomical models are currently available for some sub-skills training but do not offer the flexibility and potential of computer-generated modelling. Limitations are mainly bound by physical technologies, programming resources and the integration (uptake) of such systems within the surgical training philosophy. Since integration would potentially improve uptake it is this approach that is adopted for the design of embedded simulation in VOEU. This incurs the cost with regard to facilities and resources, particularly staff, and the opportunity costs of both trainees and trainers being temporarily unavailable for service commitment.

2.3 Development of a collective self learning approach

The aim is for real users to design and build simulation systems based upon their experience so that they teach the relevant sub-skills in context. These systems will need to integrate with regulatory authorities such as the General Medical Council, The Royal College of Surgeons of England and another administrative body which has been established termed the *Postgraduate Medical Education and Training Board (PMETB).* Self-learning will also allow us to collectively tap the wealth of surgical experience originally obtained through apprenticeship. Clearly defined learning outcomes are vital.

Users also require a suitable framework for the embodiment of surgical knowledge. Ultimately an adaptive user profile will offer freedom to explore and pursue intellectual curiosity. This has to be tempered by accommodating the moving target of the syllabus and the curriculum. The inevitable modification of these will be well supported by adaptive hypermedia (AH).

For the surgeon, skills in ICT are becoming core skills (8) and to ensure that these skills are not the 'Achilles' heal' of the research an online book has been provided that introduces the users to the key concepts in advance of using the technologies being developed (see Appendix 2.1) This encourages the adoption of this new way of working. It provides the developmental principles for end-users to create relevant material.

The potential educational value should drive the use of simulation media. A set of goals should always be defined, the instructional design prepared, and the pedagogy should specify the requirements for the simulation. Competition for resources will improve efficiency, as will the pride of users who develop components, since preparation of material satisfies an educational role in its own right.

2.3.1 Routes for delivery of a user-generated bottom-up approach

The earlier work using an *Asynchronous Transfer Mode (ATM*) training system needs to be reviewed in the context of the various modes of synchronous and asynchronous web-based training now available. Models can be integrated into simulator-based training systems, which are far more economical, widespread and better supported. With time, technologies are superseded and the process of delivery of such systems moves on.

This adds weight to the concept of a fully integrated surgical training system with optimised client server architecture that is flexible enough to accommodate such changes. This approach needs network administrators to ensure that service provision is possible. Administrators can more readily adapt to changes in technologies in different layers of information delivery. The user-generated content philosophy must however be incorporated within the already established surgical training infrastructure and represent a facet of the move toward web services. Computing is only truly pervasive if it pervades every channel open for communication with the user, not just ubiquitous access.

There is a need to establish the dynamics of the training opportunity and adapt the tools to this. This raises the issue of how to simplify the tasks that the end-user would need to perform for the data collection and integration of their own peer-reviewed material. Development in the SIM3 discussed later in chapter 6 shows how high quality images and video can be the basis for model generation. As with the answering of most questions, new ones arise, for example *'Just-in-time'* education, how much surgeons should prepare for operations, would pre-operative planning be improved by such systems if these models were available, and if they were would they be a help or hindrance?

2.4 Building a process for content development

Construction of environments from live surgical video, with operator-controlled data collection (rather than robotics), was employed recording spatial geometry data with the images. Preparation of this thesis has concentrated upon development and integration by a small group of experts. All of this is only possible due to their valuable time to support the design and validation process. Whilst implementation of user-generated material is potentially possible, currently this is not practical without end-user training. Defining how they may benefit educationally provides a basis for development.

2.4.1 Without costing the earth?

One advantage of these systems will be to minimise travel costs for the users, reducing the number of trips to national training centres. It is also possible to optimise the trips that individuals make by building the main educational components within *Multimedia Educational Orthopaedic Modules (MEOMs)* so that they are adequately prepared for the material that they will use during the courses. They may also use the course modules in clinical practice afterward by referencing back to the surgical logbooks. This will link to their individual learning agreements, set up with their trainers. Medical informatics epitomises society's march toward one potential view of the future. To set such an *'information society'* revolution in the context of social change, a historical example may serve to convey its potential impact.

2.4.2 Historical comparison with the development of simulation in aviation

The evaluation process is an integral part of development. This will include the integration of such an evaluation methodology into the iterative design process. A classic example of this is the specification for flight simulation that was referred to in Chapter 1, relating the tool to the training task. An earlier example was the Link Simulator (Figure 2.1), successfully used for basic flying skills training. More recently systems have been developed for complete training for type rating upon specific aircraft. A similar process of evolution and therefore progression from discrete sub-skills to fully integrated systems is anticipated in surgery.

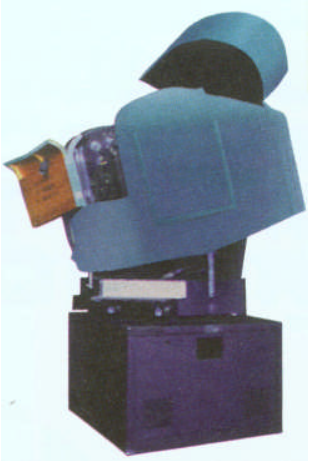


Figure 2.1 'Link' Flight Training Simulator – 1928, Edwin A. Link develops the Link Trainer flight simulator as a means of providing affordable flight training to pilots without having to leave the ground(9).

2.5 Alternative technologies for skills development

Performance assessment can only come from objective measurements of function. This requires a reliable repeatable method of evaluation of tasks. There are already a number of options available for skill assessment. Traditionally, those used included patients, cadavers, animal models and synthetic models.

2.5.1 Patients

There are ethical questions arising from teaching technical skills whilst operating on patients. Clearly, in teaching hospitals, it is not surprising that the training surgeons are, whilst under the appropriate level of supervision, performing most of the surgery. Consent for assessment of technical skill is difficult, since this implies the risk of lack of expertise in clinical practice.

There is no more valid a test than a real situation. There are questions of bias, true objectivity and the limitation of standardisation. It is not possible for an operation to both be standard and realistic. To achieve regular assessment of technical competence, it is necessary also for evaluation to be performed in other environments that can be standardised. This should be complementary to those evaluations, which take place in operating theatres. The development of an educational framework support with individual learning agreements needs to be built into the process, affording users a clear record of progress as discussed in Chapter 7.

2.5.2 Cadavers

Whilst cadaveric material provides a 3D special relationship with the body when operating, it is not dynamic. There are issues to be considered surrounding the use of such material. Important issues are that cadaveric material cannot be re-used, that chemical agents such as formalin are required for fixation and that it is not possible to use this for training certain procedures such as laparoscopic surgery. For these and other anatomical reasons, such as variability of anatomy, cadavers are not appropriate for regular skills evaluation.

2.5.3 Animal models

Use of live animal models for surgical training in the UK is not permitted, unlike the United States and on the Continent of Europe. Whereas live animals offer the advantages of tactile validity, there are limitations. These include ethical issues and anatomical issues. Higher levels of animals are required to maximise the degree of transfer of skills learned from animals to operating on humans, with comparative anatomy. Animals are expensive to maintain, requiring specialised facilities which can be under threat of violent action from extreme groups, a problem not experienced when training on humans! *Stotter, S, L.* (10) produced freeze-dried animal tissue for bench-training of surgical skills with the developing anastomosis workshops (11) incorporating bench models into the training regime. This was an early example of ‘skills stations’ with clearly defined educational tasks.

2.5.4 Synthetic models

These have improved considerably over the years. They are still somewhat low fidelity and expensive. Individual models are typically over £500 each. They often contain re-usable components, which require replacement after each session typically at £40 - £50 each. It is this cost, and also the assertion that the models do not represent anatomy well with respect to geometry and tactile properties, that has led to the need to explore computer-based simulations. At the level of basic surgical training, these may have a role for demonstration of core skills, such as triangulation of instruments and minimal access surgery. Some have been developed to the level required for higher surgical training for certain sub-skills. There is little evidence for the degree of transfer of learning skills from such models to human reality.

2.5.5 Models of skill assessment

Whatever tools are used to assist the training process, procedural lists are likely to be required as a record of the technical skills attained. This is explored in more detail in Chapter 7 when the role of the surgical logbook is reviewed in context. Such lists are surrogate for assessment of technical skill. They are ‘second hand’ analysis, reflecting the exposure to cases offered by the working environment. It is a measure of quantity and not quality. Training must be competency based. This implies that, under supervision, trainees will not be allowed to proceed to practising techniques beyond their abilities without supervision. Decreased level of supervision indicates that there is an increasing level of competence.

Some might argue (12) that it is only such *in-training* evaluations that provide indicators of performance in real world situations. This is free of the biases that result from discrete examination episodes and this comprehensive system of day-to-day assessment should be the focus of effort. The approach of the Surgical Education Group at the University of Toronto to develop a bench model examination for surgical skills (referred to as the Objective Structured Assessment of Technical Skills - OSATS) (13) suggests that certain elements can be extracted from the procedures. Typical examinations involve rotating through 8 x 15 minute stations performing a technical task. They are marked using two scoring systems. The first is a task-specific checklist. The scoring rubric delineates whether a trainee has or has not performed an element of the procedure. The second is a global rating considering surgical behaviour. These are now administered yearly in Toronto and cover such issues as performing bowel anastomoses. Of interest to the developing shoulder simulation system is the global rating scale for operative performance, which is based upon performance irrespective of the level of previous training. This covers the following areas:

- Respect for tissue, knowledge of instruments and their handling
- Time and motion, the use of assistance

- Knowledge of specific procedure and flow of the operation
- Subjective appraisal of the overall performance and the final product quality

The author acknowledges that such systems are labour intensive and expensive. However, they demonstrate a degree of validity with inter-station reliabilities of 0.8 being reported, adequate for making high-level decisions upon the trainee's ability. Usually, there is a role for such systems and these have already been well established at the basic surgical training level. At the higher surgical training level, individual stations will continue to be integrated within the overall training infrastructure. Simulations may provide HST skills stations focusing upon the special sub-skills.

2.6 Prior research into surgical simulation

This section outlines the historical perspective that has preceded the research into the development of computer-based surgical simulation presented in this thesis. The review of the field outlines the early experiments and classical approaches to so-called surgical simulation in the era of pre-digital technology. Tools such as the anatomical models were originally developed in the seventeenth century[3]. These were often mannequins, made of leather, representing the foetus. They were used to train midwives in the art of assisting childbirth. Those made in France were reputed to have been of the finest quality. Here is an early example of a clearly defined training scenario for which the 'simulation' has been developed. Whilst plastic models exist today for arthroscopy training, these lack realism and are costly due to the need to replace disposable components.

2.6.1 How digital simulation might benefit from alternative technologies

The ultimate aim is for digital simulations to pass a 'believability test'. Broadly, this means that it is possible for users to interact in such a way with the virtual environment that they would be *unable to distinguish between the virtual environment and the real one*. This is also referred to as an A/B comparison, comparing the two training scenarios, the simulation and the reality. This leads to a sense of presence. Integrating digital technologies into simulated training environments may not be achieved through computer simulation alone. Rather it may also be important to

---

[3] Also of note were the wax models for the intricate anatomical detail, of which few remain. See the Hunterian Museum at the Royal College of Surgeons of England (www.rcseng.ac.uk) in London.

consider other aspects of the simulation studies such as the draping of the "virtual patients", the gowning and masking of the operator, and a similar physical environment to the operating theatre.

To minimise unnecessary financial expenditure, compromise has to be made to limit both the valuable resources of human time and disposable equipment. In the early experiments with the simulator it was considered essential that a draped mannequin was used to provide the shape of the shoulder in order to at least produce some semblance of a physical simulation of the operating environment. This principle of the educational environment has been overlooked in the case of most of the technologies outlined below. Much is gained by establishing a simple training environment, but the cost of more sophisticated operating theatre environments may be difficult to justify for training.

2.6.2 Prior research on surgical simulation using computer simulated virtual environments

Development of surgical simulation is based upon a regional demand. There are several centres worldwide which have led the field. Naturally the developments tend to reflect the state of industrialisation of countries with much of the work being carried out in the US and Japan. The following main centres have been singled out for attention due to the specific skills and expertise that they demonstrate, starting in Europe.

2.6.3 Germany - Forschungszentrum – Karlsruhe (FZK)[4]

The FZK has been developing the field of deformable organ modelling *(Figure 2-2).* This team adopts a high-end approach, particularly with respect to obtaining high-resolution haptic feedback with three kilohertz (3 KHz) frequency sensors for boundary definition, rather than using a vector-based modelling approach, as seen in the Prodiscus System (page 18 below). Originally developed as an offshoot from the nuclear decommissioning industry, these technologies again relate closely to those employed in robotic surgery (Figure 2.3) but are outside the scope of this thesis.

[4] More information is available at: http://www.iai.fzk.de/englisch/default.htm

Figure 2.2 FZK minimally invasive surgery simulator (2001)

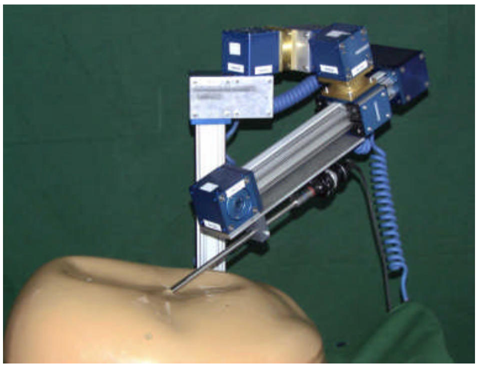


Figure 2.3 FZK Telerobotic system for controlling surgical instruments

2.6.4 Germany - Helmholz Institute Aachen (HIA)

The hip simulator developed by the HIA focuses upon open surgical technique training for the development of periacetabular osteotomy skills (*Figure 2.4*). This is complex hip surgery likely to be performed by only a small number of surgeons. It does however demonstrate the principles of anatomical and biomechanical modelling and the role that simulators can play in both the learning of surgical technique and pre-operative planning. The segmentation approach has medical applications elsewhere (14).

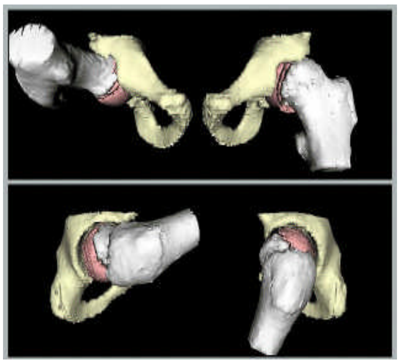


Figure 2.4 HIA Hip simulator (displaying segmented anatomy)

2.6.5 Italy - Scuolo Superiore Sant' Anna (SSSA)

This simulator focuses upon the knee and mechatronic tools development. It is targeted for the basic training of surgeons *(BST)* in the domain of arthroscopic surgery, concentrating upon geometry and mechatronic tools rather than high-end visual simulation at this stage (Figure 2.5). A UNIX platform is used, again with biomechanical and haptic rendering of the wire frame model.

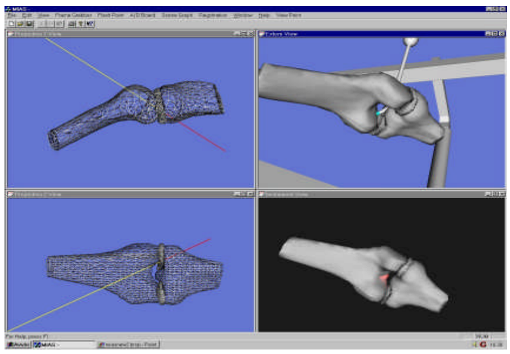


Figure 2.5 Arthroscopic Knee Simulator from SSSA demonstrating the wire frame rendering

2.6.6 France – Grenoble

This *sacroiliac screw fixation simulator* is specifically designed to train surgeons in the use of certain tools. These are surgical simulators for Image Guided Orthopaedic Surgery (IGOS) Systems (Figure 2.6). The user population consists of expert surgeons (surgeons that know the traditional procedure) who want to learn Computer Aided Surgery (CAS) procedures.

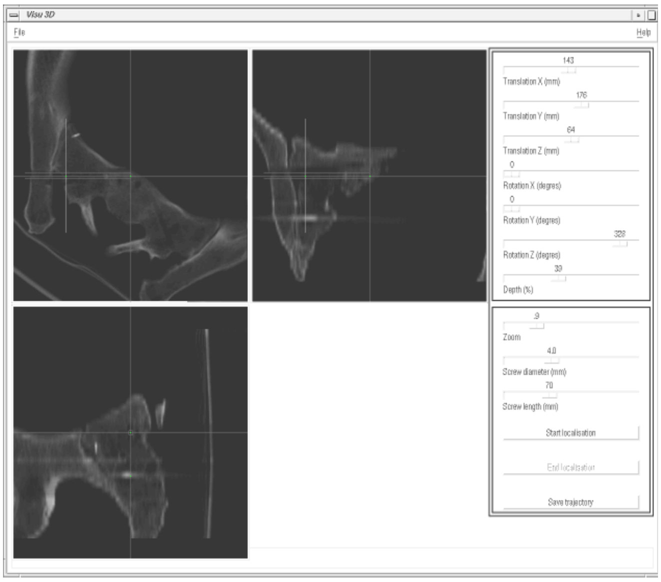


Figure 2.6 Grenoble Simulator for sacroiliac screw fixation using a CAS system.

2.6.7 Sweden – Prodiscus Shoulder Simulator

This company markets many simulator systems derived from technologies developed as part of the motor/defence industry in Sweden, and bought under licence. Two are important here, the shoulder simulator (Figure 2.7a - near right), and the Minimally Invasive Surgical Trainer (MIST) VR psychomotor co-ordination simulator (Figure 2.7b – far right)

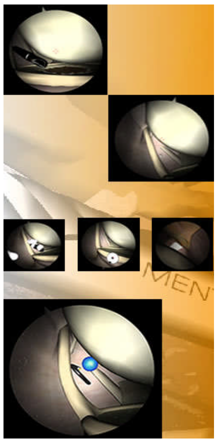


Figure 2.7a Prodiscus Shoulder Simulator

Figure 2.7b Prodiscus MIST simulation (originally developed in the UK and now sold under licence.)

2.6.8 United Kingdom

Examples of simulators, such as the MIST VR referred to above (Figure 2.8), demonstrate the interplay between political and commercial expediency versus academic credibility. Both have a significant role to play in the development of innovative systems. With simulator systems taking 5 - 10 years to develop, the option of developing sub-skill simulation systems, such as MIST VR, are certainly commercially viable and therefore attractive. These also offer the opportunity to provide basic building blocks for future development.

2.6.9 United Kingdom - Exeter Virtual Worlds Simulator

This was the simulation used to demonstrate the principles in the early experimental trials. SIM1 was a product of the 1st prototype detailed below. See Figure 2.8.

Figure 2.8 Early demonstration of EVW simulator(1996)

2.6.10 United Kingdom - Sheffield

This research experiment (15) did not develop further than an early prototype (Figure 2.9). It exemplifies many of the early examples of simulation where the hardware was still prohibitively expensive for the development of realistic prototypes and hence, in isolation, is unlikely to produce sensible business models for their future development. The work continues with EPSRC funding.

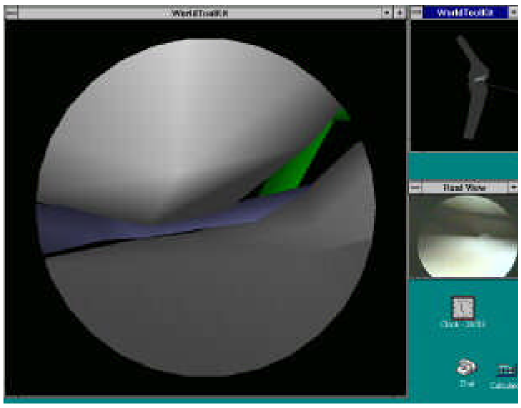


Figure 2.9 Sheffield Knee arthroscopy simulator

2.6.11 United Kingdom – London

The MIST VR (Minimally Invasive Surgical Trainer) has been used to assess a stand-alone desktop system that is designed to train a predominantly visual skill. This perception is reinforced by the assertion that the relative importance ascribed to each sensory modality can be compared when performing the necessary skills of arthroscopy.

Hand-eye co-ordination, triangulation and tissue manipulation can be practised on such relatively simple simulators. Originally the MIST VR simulator (16;17) comprised a 200 MHz Pentium PC with 32Mb RAM linked to a jig, which had two "laparoscopic instruments" in gimbals, attached to motion-detecting potentiometers. This allowed the trainee to grasp a target, pass it to a second instrument, and place it in a randomly positioned cage. Further tasks such as applying diathermy to nodes, which appear on the ball, can be practised. An attempt has been made to define metrics for this (18). Prof. Sir Ara Darzi has reported (19) advances in the system.

The advantage of low 'cost' technology systems, such as MIST VR, is that they allow training in basic endoscopic skills such as hand-eye coordination, triangulation and manipulation using relatively inexpensive computer platforms. The disadvantage of the earlier version is, of course, that such a basic system is non-immersive and it bears no relationship to real surgery performed in a real patient, for there is no 'Surrounding World'. The virtual environment (VE) is abstract. It is no good training a surgeon to ligate a duct (tie off a tube) if the trainee cannot differentiate between the cystic duct and the common bile duct. The former cures, the latter kills. Its value lies in psychomotor skills training rather than pattern recognition training.

2.6.12 United States of America

In the latter part of the 1980's, centres such as the University of North Carolina and the US Department of Defence worked upon the concept of future surgery, using head mounted displays for VR procedures in surgery, mainly relying on adaptation of military software and hardware. By 1995, Col. Richard Satava proposed categories of achievable VR applications for medical and surgical domains (20). These covered the fields of surgical planning, therapy, prevention, training, skill enhancement and data base visualisation. Sponsored by the *Defense Advanced Research and Procurement Agency (DARPA),* this concentrated upon robotic and telepresent surgical systems using VR. This was related to the fact that, at the time, the perceived need was for ways of dealing with nuclear, chemical and biological weapons and that telesurgery might be one way of handling battlefield casualties. Systems did, however, and then start to evolve for the rehearsal and planning of other surgical procedures such as total joint replacements and ophthalmic procedures. This is combined with the concept of augmented reality where 3-D graphic projection is brought into the operative field. Higgins and Satava at the University of California San Diego (UCSD) obtained funding via the US defence department for simulation development, reputed to be $6 million in 1997. This has been developed for a clinical skills laboratory in San Diego (21).

Further enhancement came in 1993 when magnetic resonance images (MRI) were first overlaid onto a near real-time video image of a human head (22). Indeed, some of the early pioneers such as Tony DiGioia (23) are still persisting with augmented reality in robotic systems and developing the domain of Computer Assisted Minimally Invasive Surgery (CAMIS). This work, though fascinating and likely to integrate well with the theme of this thesis, is beyond its scope.

There has been commercial platform development, with *High Techs* Plantations (HT Medical) and *Cinemed*, attempting to exploit the earlier technologies (24) and up to date records of progress are kept both on the relevant websites (25;26) and in publications (27). Probably the greatest core resource to be developed in the United States in the last 10 years is that of the *Visible Human Project (VHP)* (28), which provides data sets for the construction of three dimensional models from microtomed human bodies of consenting executed prisoners. There are both male and female specimens, of which the female can provide an isotropic data set and normal visual resolutions, since the slices are 1/3 mm thickness, providing a 75 dot per inch resolution. The purpose of development over the next ten years is to see digital integration with the dataset using patient-specific data scans (such as CT or MRI) to construct 'near real time' viewers of the data for individualised preoperative planning and training exercises. Already the anatomical models have been generated and are viewable over a network, commercially developed from Primal Pictures (www.primal.com).

Other simulation systems are being developed, such as at the University of California Berkley VESTA Project (Virtual Environments for Surgical Training and Augmentation (29)), which tend to take the virtual human model one stage further by including dynamic, tactile, visual and auditory modes. Whilst such systems are at the forefront of VR, they are still dependent upon high-end computing resources to manage the graphics and special modelling, and thus currently do not provide suitable solutions for training. They provide high fidelity libraries with virtual anatomical and physiological data sets. At present these libraries would be unlikely to be supported by in-house resources with the current infrastructure where the end-user is situated. This matter of being able to support such high-end resources in an ICT environment with the rapid onset of technological redundancy is also addressed in this thesis.

2.6.13 Japan

Progress in Japan is difficult to assess since most of the work appears to be in the commercial sector with companies such as Hitachi (hearsay), and therefore not accessible. Only public domain information has been used in the preparation of this thesis.

2.6.14 Current developments of other centres worldwide

The progressive development of the field with both academic and commercial development means that the advances in computer resources will be passed on to working prototypes, but may or may not be announced publicly. The best ongoing database, to which the author contributes, is held at SSSA[5], developed as part of the VOEU project.

These different approaches to simulation systems lead to the concept of scenario classification for simulation. A simulator may be for planning or interventional training, and represent an aspect of image guided or open surgery. To this end it is possible to classify the systems according to a simple matrix, as outlined in Table 2.1: *NB*; No available systems at the time of writing are able to model open surgery (Immersive 3D view).

[5] Scuolo Superiore Sant' Anna (SSSA) database is at: http://www.sssa.it /

Feature	**Development**	**Image Guided Orthopaedic Surgery (2D view)**
Autonomy	Mixed Modal Complete	**EVW,** HIA, Grenoble, Sheffield FZK, SSSA
Interactivity	None Single Mode Multimodal VR feedback	 Grenoble, **EVW,** Sheffield FZK, US, HIA, SSSA
Presence	Abstract Conceptual Realistic	MIST HIA, Sheffield, FZK, SSSA, Grenoble **EVW**
Environment	Stand alone Embedded Integrated	FZK, Sheffield HIA, Prodiscus, SSSA, Grenoble **EVW**
Scenario	Time frame Pre-operative Intra-operative *Generic* *Specific*	HIA, Grenoble FZK, SSSA, **EVW,** Prodiscus Sheffield, Grenoble *MIST* *All others*

Table 2.1 Simulation system features

2.7 Developing a new orthopaedic educational strategy

The heart of developing a new educational strategy is the *Instructional Design (ID).* Relevant to surgeons and computer scientists is the assertion that the developing multimedia integration strategy must be incorporated inside the pedagogy so that it accommodates the syllabus for specific populations of users preparing for certification examinations, such as the *Certificate of Completion of Specialist Training (CCST).* This pedagogy must be adapted to the curriculum to take the resources and opportunities available to each individual into consideration. *Adaptive hypermedia* is employed to support this required adaptability.

The evaluation process, designed to ‘close the audit loop’, is seen in the context of developing user understanding as well as the ability to apply the knowledge in accordance with Bloom’s Taxonomy (30;31).

2.7.1 The rationale for technical skills competence assessment

The clinical environment requires a broad range of skills. The criteria for the ideal surgeon have been identified by the Royal College of Surgeons of England as consisting of the following:

- Communication skills.
- Knowledge of basic sciences.
- Knowledge of theoretical clinical skills.
- Knowledge of clinical skills.
- Decision-making - treatment options.
- Surgical skills and manual dexterity.
- Postoperative management.
- Teaching and learning skills.
- Management and leadership skills.
- Research and data analysis skills.

Spencer remarked that a "skilfully performed task comprises 75% decision making and 25% manual dexterity" (32). It is likely, however, that the 25% is the largest individual component of the surgeon's expertise, which is not assumed by other specialities. Therefore a rationale is needed for technical skills assessment. There are 3 main categories of assessment for surgical performance training. These are:

- Formative evaluation
- Summative evaluation
- Evaluation of continued competence

These refer to evaluation of procedural and declarative learning. Formal independent testing is fundamental to the process of considering medical education outcome and its re-validation. Assessment methods for each of these may be different. The formative assessment method implies an ongoing method with feedback given to trainees. This can consist of proficiency, refresher and "just-in-time" training. A relatively informal system of assessment is discussed in Chapter 9, in the context to the Joint Committee for Higher Surgical Training structured system for appraisal and assessment. The option though is for this to be more formal as proposed by Kopta (33). He suggested that, consequent upon reducing the technical evaluation to specific components, this allows for concrete assessment of technical performance. This appears to be borne out by the current approach of an ergonomic task analysis, which then forms the basis of an evaluation procedure. Summative evaluation implies that the final grade or decision upon an aspect of performance is reached. It also implies external recognition of such achievement. The proposed reconfiguration of surgical training to distinguish between higher surgical training and specialist training is likely to comply with certification or licensing, such as the CCST, and needs support for its infrastructure as there is a general move toward competency-based training rather than experience-based training.

2.7.2 Maintenance of competence

Skill bases degrade with time, so-called "skill fade". The question of currency thus arises. The simulation system is currently not designed for *continuing medical education (CME)* applications although it is likely that CME systems will ultimately be introduced as a way of regulating the profession. It is possible that, through the system of clinical governance, surgical outcome audits may highlight certain individuals who will need to demonstrate that they have maintained their technical competence, or else a process of re-validation, on a 5 yearly system. This may be introduced in the UK depending upon *General Medical Council (GMC)* recommendations. This process will become more focused with the growing atmosphere of public expectation that the medical profession will improve its ability to ensure the competency of its members following certain high-profile cases, such as the Bristol children's heart surgery service (34).

2.8 Influences upon the choice of surgical simulation prototype

Computer-based simulation is simply the application of ICT to the old problem of how to convey the relevant information and skills to trainees with the limited resources available. Media limitations have restricted the development of surgical training. This should not necessarily limit future plans for training. Shoulder arthroscopy has been employed clinically for the last two decades. There has been a spectacular growth in the technique (35). This has been chosen as the focus for vertical prototype development. As a relatively new technique, where trainees have little opportunity to experience the real surgery, this offers trainees the chance of gaining experience in a low risk environment, that is predominantly used for procedural training.

There is a need to develop computer-based surgical simulators (36-38). A surgeon navigating the shoulder joint through manipulation of the arthroscope has to accommodate a view of anatomy as seen (magnified) through the eye of an arthroscope. The surgeon performs skills such as locating structures and tracking between them in an orderly way. Currently the visual complexity of the shoulder joint makes unrealistic demands upon a networked graphics-based approach, which remains impractical for high-fidelity pattern recognition training. The educational environmental pressures directed our team toward the video based approach.

2.9 Overview of shoulder arthroscopy

The indications for shoulder arthroscopy are varied. Most orthopaedic procedures (musculoskeletal surgery) are for the relief of pain or the improvement of the range of movement. The development of minimal access *'keyhole'*[6] techniques has arisen from the intention to supersede the open surgical techniques that generally causes greater scarring and requires more rehabilitation. The profession is now exploring ways of passing on these newer techniques to surgeons in training (Figure 2.10, displaying the normal operating environment).

Figure 2.10 Shoulder arthroscopy set-up in the operating theatre

The skills to be learned and assessed require the development of a procedural training tool employing multiple visualisation techniques, from conventional stills and

[6] Minimal Access Surgery (MAS) is also referred to as Minimally Invasive Surgery (MIS), or 'keyhole' surgery in lay terms.

video to complex virtual environments (VEs). The virtual environment should allow orientation, navigation and pattern recognition in the three dimensional virtual environment to mimic actual surgery. The simulation system depends on a suitable system of distribution to meet the key specifications of real-time image display and low latency feedback.

2.9.1 Justification for investigating shoulder arthroscopy using video-based simulation

An alternative to near-real time graphics-generated VE models considered for training in shoulder arthroscopy is visual feedback using video. During arthroscopy the picture on the monitor is often recorded on videotape, and occasionally directly to MPEG[7] files. This is to assist the surgeon in diagnosis, or for later review with colleagues, or indeed for tutorial-style, rather than skill-based, training purposes.

The work underpinning this thesis suggests that *sufficient video images can be stored to provide appropriate visual feedback to the surgeon practising on a model representing the shoulder*. One can then simulate the movements of the arthroscope, which determine the appropriate video images on the display monitor. The surgeon can thus interact in an environment providing a high degree of realistic manipulative and visual feedback.

Learning arthroscopic skills requires the ability to navigate, through recognition of both anatomy and pathology. The potential difficulty with this approach arises when one attempts to modify the environment. With training in specific procedures such as repairing dislocations and stitching rotator cuff muscles (the muscles around the shoulder joint), such modelling of surgical procedures requires the incorporation of a temporal dimension *(t)*. This is *forward only* for the purposes of this work. There is a logical forward progression, a sequence of events. The order is vital. A scene cannot be displayed that is out of sequence and still expect the operator to maintain a sense of presence in this virtual environment.

Complete immersive simulation is the ultimate goal. The starting point is to develop a system where it is possible to simulate the whole of the 2-Degree of freedom (2 *dof*) mapping of the 4-Degree of freedom (4 *dof*) environment in which the surgeon

---

[7] Facilities are available in Royal Bournemouth Hospital for recording in MPEG1 and MPEG2 formats (Smith & Nephew Ltd.).

would normally navigate. This is described in detail in Chapter 3. As the integration of the simulation evolved, this demanded that the integration and simulations employed became simpler. The price of this is loss of immersion.

2.10 Surgical users requirements for simulation validation

Having built the models to integrate with the instructional design, to make such simulations valid educational tools, surgeons will need to use the simulations for their future evaluation. It is vital to achieve a reliable gauge of performance from the simulations before they can be ascribed such a role. The analogy is drawn with the method of formally assessing vehicle simulators, such as commercial flight simulators. This approach to education has been successful for over thirty years. The introduction of new JAR STD 1α guidelines from the Joint Aviation Authority (JAA) set standards for such simulators in the field of aviation. Other industries (*e.g.* space exploration and the nuclear industry) have invested heavily in the use of simulation for skills training, but the main driving force has been military vehicle simulation.

Since the evaluation of simulators is going to be linked with specific tasks and scenarios, the starting point remains the task which may benefit by simulation in training. The quality of the representation of the tasks in the VE depends upon quality of the simulation and thus will affect their educational value. They are:

- Simulating critical events so that there is no risk to the patient.
- Reproducing a specific critical event repetitively.
- Providing an opportunity to hold simulation for discussion and teaching.
- Allowing errors to be made and explanation of their consequences given.
- Provide a record and critique of performance.
- Allow objective evaluation of performance.
- Allow uncommon events to be experienced.
- Allow control of independent multiple variables.

Simulation may benefit surgery greatly providing both assessments of competence and evidence of retention of expertise. However, other aspects of surgical management may only be approximately modelled. Educational scenario models have a role to play in the evaluation of simulators.

2.11 Summary of the overview of surgical simulation

The envelope of exploration for this thesis is the educational system into which such new methods must integrate the technological limitations of the day and the user, who may be obliged to adopt such systems. The first chapter indicated the need for a review of the surgical training, outlining the scope for change in Orthopaedic Educational Infrastructures. This chapter described the ambition to build the tools to manage educational content and performance assessment in a standardized format that is independent of the content itself, but interchangeable between parts of the integrated system. It should be possible to evolve the educating and learning processes into true life-long learning with technologies supporting an individual's self-directed *continuing professional development (CPD).* Any future surgical training system is likely to include the adoption of the credit system used in CPD and would allow for the flexibility of simulation integration.

Instructional design (ID) will need to accommodate all stages of design from teaching or learning activities in order to accomplish stated objectives. These will be criterion-referenced measurements of learning outcomes. As part of the process, course convenors will need to assess learning needs (goals, constraints, priorities). The Virtual University described in Chapter 7 provides an educational perspective and toolkit driven by this philosophy. Simulators alone are not the answer. If such simulation is to be used, then reliable measures of performance need to be defined and the simulation validated. They will also need to evolve Course Design Methodologies that will employ recognised structured surgical course models. These can be adapted to *.xml* as described in Chapter 7.

Future surgical education systems will have to address *Curriculum Design* issues (39). For this a curriculum strategy is required and *The Royal College of Surgeons of England* is currently developing this for *Higher Surgical Trainees (HSTs).* Ultimately for surgical courses, this will probably still be content driven, since exposure to surgical experience is so vital to understanding and practice. Such simulations may however be included for the training of certain sub-skills initially.

Examples of technology-supported learning applications include computer-based training systems, interactive learning environments, intelligent computer-aided instruction systems, distance learning systems, web-based learning systems, and collaborative learning environments. Whilst attempting to combine many of these principles, the future development of surgical simulation may depend upon the ability to build such tools into the curriculum, the *caveat* being their proof of effectiveness. The ultimate aim is to develop an *outcome driven* system delivered in a safe environment which though *technology dependent* is not *technology driven*.

2.11.1 Summary of educational factors relating to distributed simulation in healthcare

To the question of how to best pass on surgical skills from generation to generation, words are not the answer (40). With progression up the ‘imaging elevator’, the hierarchy of imaging, we now stand at the watershed of the next major evolution in training methodology. To navigate this new medical world order, a paradigm shift is occurring culturally, technologically, and socially. Thus exploration is needed of the development of the relationship between hypermedia and the future role of pedagogical toolkits for the development of adaptive strategies. The educational system demonstrates a scalable infrastructure that accommodates the individual user’s requirements.

The mastering of limitations of the *physical world environment* in vehicle training suggests that the anatomical limitations in surgery can also be overcome but possibly the cost-effectiveness and educational value will be harder to prove. The variables that are important have been defined, and the research envelope bounded.

Users require reliability and accessibility and so by developing an experimental prototype to explore the constraints (*e.g.* educational outcomes) it was possible to create a framework adaptable by users to meet their needs.

In conclusion, it should be possible to improve surgical training by the introduction of complex video-based hypermedia, embedded within an integrated educational environment, though to achieve this will require multidisciplinary research backed by the established institutions. This will have far reaching implications for educational change particularly at the individual and regional levels. The next chapter explores the design and development of a simulator for shoulder surgery.

3 Shoulder simulator design and development

3.1 Overview

This chapter establishes the area of surgical technical knowledge from which the thesis explores integration of simulation into surgical training. The thesis focuses more upon the human computer interface than the databases themselves. It includes qualification of the appropriate physiological thresholds of performance. These impact upon system design. The basic physiological targets that human performance limitations set, and thus the thresholds of performance that VEs need to set to appear real to the user, are defined.

3.2 Presence of mind

In order to perceive the image as moving, with no flicker, a frame refresh rate greater than 20 per second (Hertz - Hz) is required. This can be performed in real time or put together as an animation sequence and then replayed. For *minimally invasive surgery (MIS)* to work, near real time is required. One needs to achieve *"The willing suspense of disbelief"* in the mind of the operator. This term, poached from acting, is used to describe the psychological transformation necessary to induce the optimal learning environments, where the individual temporarily ceases to perceive the procedure as simply being a part of their training schedule. In VR modelling this is termed *Presence*. There are some fundamental reasons for adopting this approach: *e.g.*, operating on real patients carries a significant potential risk related to the indications[8] and complications of their arthroscopic shoulder surgery. To a lesser degree, there are also risks for the operator. This may be a physical threat to the self in the form of an

[8] Indications and complications of shoulder arthroscopy are outlined under the appropriate headings of the shoulder section in VOEU under http://voeu.ecs.soton.ac.uk/VOEU/library/ section marked: VOEU : Library : Shoulder : Arthroscopy . This is a virtual path rather than absolute and so needs to be located via the main VOEU site.

inoculation of a viral pathogen such as hepatitis B, C or HIV, or a legal complication as a result of malpractice litigation.

The associated *Multimedia Educational Orthopaedic Modules (MEOMs)* could accommodate this aspect into the simulation environment by incorporating relevant sections *e.g.* microbiological cognitive recall as one of the sub-skills requisite in the pre-operative section.

So how can this 'panacea' for surgical training be achieved? If virtual reality modelling is considered in the traditional sense, the trainee should be induced to perceive the risk of consequence associated with failure to achieve an acceptable standard of performance, where in fact no actual physical risk exists.

The particular training environment in which the trainee is learning and demonstrating their skills can generate this. For example the trainee could be using such a simulator as part of the regular professionally supervised training schedule, *e.g. Continuing Professional Development (CPD)* or part of their 'specialist' training scheme *(e.g. Higher Surgical Training - HST, in the UK)*. The demands of the examinations, standards for qualification and ultimately of the work practice focus upon hurdles such as the *Certificate of Completion of Specialist Training (CCST)* and preparation for this is integrated with the effort an individual makes during their specialist training programme. They are subject to the *Regional In-service Training Assessments (RITAs)*, a process for monitoring the progress of an individual. Then the validity of their performance would necessitate their attempting to achieve the best possible results.

This thesis describes the development of a method for the evaluation of the performance of the arthroscopic simulator. Since the human mind is adept at "filling in the blanks", users will overcome the shortfalls in certain areas of the simulation, such as the interactivity to control the fluid irrigation normally used in arthroscopy. The limitations become less significant to the trainee if they are receiving adequate simulation via other senses. The principle was first demonstrated in mechanical simulators, such as the 1920's Link flight simulator see Figure 2.1 and the motorcycle simulator that provided a representative sensory experience of riding a motorcycle known as the *Sensorama* built by Morton Heilig in the 1960's (41).

In a surgical operating environment there are certain aspects of reality which people use, such as types of clothing and drapes to demarcate who is able to perform tasks in the sterile arena and thus able to work within the operating field, or the recognition of other people within the surgical team as fulfilling discrete designated roles *e.g.* nursing staff, operating department assistants and the anaesthetist.

Teamwork is essential. This is often overlooked. In any real operating environment this is a significant factor in the smooth running of a procedure, both with regards to safety and the satisfaction of all those who participate. Though not fundamental within the actual simulator itself, teamwork can be emphasised within the tutorial part of a training package to achieve a compromise between an individual's role within the team and the focus upon training for skills. One of the foundations of building trusting relationships in healthcare is the ability of individuals to communicate clearly and concisely by distinct use of terms and language (42).

An intrinsic advantage of designing, building, and emulating arthroscopy (for that matter any minimal access surgery) is that the trainee obtains a monitor-based view of the operating environment, hence is already immersed within the appropriate operating environment. This offers considerable advantage since the display system interface mimics the normal colour monitor (HDTV) display with which all operators are familiar, avoiding cognitive dissonance. Surgeons can therefore focus upon specific purpose such as pattern recognition sub-skills training.

3.3 Design requirements for shoulder simulation

Ultimately society is seeking the reassurance of regular training and competency checks upon surgeons using simulation, in line with plans for revalidation akin to the aviation industry. This leads to the drive for accurate measurement of personal performance. Such measures may be *absolute*, *e.g.* advanced simulation of specific procedures, *proxy* measurements such as psychomotor co-ordination reflecting the core skills a surgeon needs to operate effectively, or the use of surgical outcome measures that are dependent upon the whole healthcare system and patient factors, rather than just the surgeon's competence, although this is a factor that has a significant impact upon outcome. Outcome really represents team performance, and it has been suggested that one of the key roles of simulation is to ensure teams can work effectively together[9].

[9] This approach has been emphasised at the Bristol simulation centre that focuses upon anaesthetic training.

The embedding of simulation into a virtual educational environment should therefore take CCST training into consideration, so that its application is relevant to the demands upon and of the end-user, who themselves are often undertaking *Just-in-time Training* on the job, making learning opportunities of their clinical experience. The plan is to incorporate this into the overall environment by developing *case-based learning (CBL)* and updating reference material based upon the individual's reading. This is detailed in Chapter 7.

3.4 Principles of human performance and its limitations

Considering the principles of *Human Performance and its Limitations (HPL),* lessons learned from the vehicle-operator training industry are applied to suggest a suitable framework for proto-standards in surgery, adapting those standards already used for commercial pilots. This section highlights the differences in tasks and responsibilities that would need to be reflected in future standards-setting procedures for surgery as compared to say, aviation. By focusing upon the physical skills that can be modelled, key areas of surgical attributes (sub-skills) for which performance parameters can be quantitatively validated are defined.

Some skills cannot be validated in this way (43;44). There is a need to integrate such training and assessment tools within a complete system which recognises and deals with other parameters that are difficult to measure using current simulation technologies. Based upon a *Human Attribute Set,* the aim is to develop databases that are both compatible and comprehensive. A subset of this is the primary attribute set. The integrated simulation system being established dictates what need be included within this. The interaction between the attribute set of the surgeon and the system (the Attribute – Performance Analysis Engine interface) is therefore an extension of the AIPES scoring system developed in Chapter 4 since it reflects the environment and scenario for training and the suitability of the simulation to mimic this. To produce a cost-effective solution to the surgical training problem, the field needs to be able to adequately justify its evaluation metrics. This reflects in part, the eventual usability of such a system.

3.4.1 Historical perspective on human-machine interfaces

In 1765 James Watt patented his version of the steam engine. He had been working as a 'mathematical instrument maker' in Scotland. One legacy of this is the dependence upon pressure gauges and dials to control systems within defined parameters for safety and efficiency. Gauges that make up the internationally accepted flight panel owe their roots to the Georgian age of steam boilers and brass. As technologies filter down to the new applications, the infrastructure for building cheap and reliable gauges was already in place by the time of the Wright brothers[10].

These dialled devices with needles designed to be read whilst both feet were placed firmly on the ground make up the bulk of engine and flight instruments. As aircraft became more complex, so the need for the number of parameters to monitor and hence the number of gauges increased. Since these need to be cross-referenced with each other, the interpretation of their measurements needs to be assimilated quickly. This process and the surrounding issues led to the vital field of aviation psychology. The reason for exploring this subject is to form the framework for surgical psychology in relation to proposed future surgical interfaces.

Figure 3.1 Integrated Cockpit 'mock-up' (flight display) for the Joint Strike Fighter

The invasion of the modern cockpit by electronic displays has therefore placed enormous demands on the engineering psychologist. In modern military and civil aircraft, multifunction cathode ray tubes (CRTs) have replaced many of the conventional instruments. The potential of these displays to present information only when it is required and in colour (for example to overlay radar returns on a moving map, and to

[10] In 1903, the initial Wright flyer had an air speed indicator and stopwatch but no specific engine instruments.

display check lists and lists of emergency actions) suggests a safer interface, less prone to cognitive dissonance. The prioritisation of tasks (aviate, navigate, communicate) should also be incorporated to optimise workload. A fundamental difference between electronic displays and traditional instruments is the capacity of the former to present and thus assist integration of widely differing information on the same display surface. The consequence of this is that unlike conventional instrumentation not all information is displayed constantly. Heavy reliance must be placed upon the computers, which decide when the pilot needs to be warned of an incipient or actual system failure. This, in turn, places an onerous load on the designer to anticipate all possible combinations of such failure.

Presenting the pilot with a well-integrated display combining complex data to provide a relatively simple and clear picture of the aircraft's situation *(Figure 3.1 above)* also means that the pilot is no longer in a good position to question the accuracy of the displays, which can lead to the phenomenon known as *over-trust*. The control and display systems of such aircraft are invariably highly computerised, making a large range of control and display options or modes available to the pilot, and several accidents have occurred due to failures of pilot's *mode awareness* (*i.e.* the pilot believed that he/she had programmed the aircraft to behave in a different way from the actual programme). The transition to glass cockpits is therefore toward information integration. This fulfils the following three tasks that are essential factors in aviation safety:

- Improve situational awareness
- Optimise workload
- Provide instrument crosschecks

Surgical interfaces should achieve equivalent results and could learn much from the human computer interfaces used in the aviation industry. Figure 3.2 reveals a computer assisted surgery display used intraoperatively. This has not been evaluated to the same degree as the aviation interfaces, but has been assessed for its psychological impact as part of Dr Troccaz’s work (45). This also led to information integration that may benefit from ontology development (46).

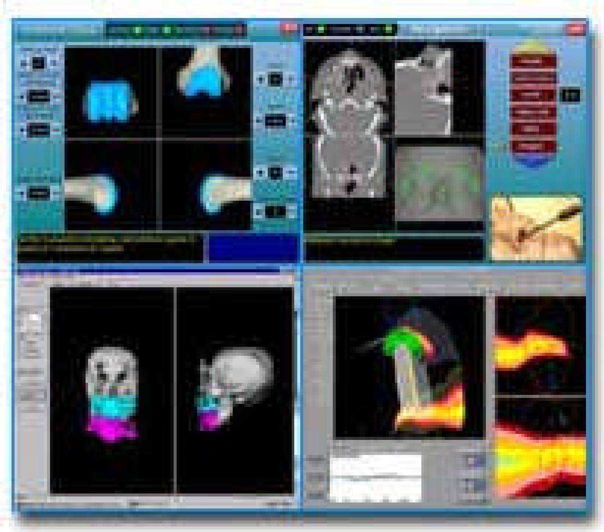


Figure 3.2 Praxim (France) surgical navigation systems interface

3.5 Human attributes that need to be modelled in surgery

This section outlines the key surgical attributes for which performance parameters can be quantitatively validated. The skills cannot be validated in this way, since (a) a complete validation system is needed capable of recognising and dealing with other parameters that are difficult to measure and (b) what is straightforward to measure is not necessarily useful, and vice versa.

The human attribute set is a list of attributes that needs to be modelled so as to develop databases. These must be both compatible and comprehensive. An accurate primary attribute set is important. The tools should be fit for purpose. The interaction between this attribute set and the system (the so-called *Attribute – Performance Analysis Engine* interface) is specific to the system being designed and the sub-skills being tested. In this case, the attributes are proficiency in orientation, navigation and pattern recognition.

3.5.1 Visual interfaces for surgery

The surgeon in their operating environment usually depends upon a good visual field. Ideally simulation would match human performance. For a human to fix on a point, the central foveal Field of view (FOV) region of the eye has the greatest concentration of light-registering receptors, designed to detect light of wavelengths that are perceived as colours. These are termed cones since this is the point that is used to look at in detail (reading text *etc.*). It therefore produces the highest resolution. There are 147000 cones per square millimetre in the fovea. This part of the eye and the associated visual mapping in the cerebral cortex limits the ideal performance for the human-computer interface. In fact, true colour requires '24 bit' data. In the Exeter Virtual Worlds (SIM1) model, a compromise of 3-colour 16-bit was adopted, allowing up to 10^6 images using 25:1 JPEG compression. The size of the image is limited by the band width and supported frame recording rate. The rate was 7 frames per second. This limited the number of frames that could be recorded in the time available using fresh cadaveric material (see page 54). Users were satisfied with the colour quality of the images. Viewing is on a high definition television monitor in operating theatres and so a display resolution of at least 72 dpi is appropriate.

3.5.2 Haptic and auditory interfaces for surgery

The haptic and auditory feedback in simulation require a relatively narrower bandwidth and so are not considered a problem for present calculations regarding data storage, needing only approximately 10% of the storage space of visual data. It is important to consider background noise levels and other distractions that can impair performance of the individual. Open surgery is often performed wearing lead jackets (γ-radiation protection) for long periods under intense lighting at a raised temperature to achieve the optimum environment for the patient, not necessarily for the surgeon. Arthroscopy is somewhat easier in this respect, not requiring lead jackets and usually not lasting more than one hour per case. Since the purpose of the simulation development is to focus upon the vision aspects for training pattern recognition skills, the detailed issues surrounding haptic and auditory feedback are not addressed here.

3.5.3 Psychology in the clinical environment

It is sensible to explore the possible methods for the introduction of simulation into clinical practice. By considering the qualities of a training environment and current proposals, this indicates the prototype training environment limitations. Surgery can be viewed as an interaction between 3 particular elements:

- The surgeon - training, expertise, and competence.
- The patient - their expectations, pathology, anatomy and physiology.
- The environment - operating theatre and hospital resources.

Decision Making: For successful surgery, elements of decision making need to be properly integrated and mutually supported. Human technical competency is probably still the most significant factor that affects outcome, because of the human capacity for sensory and perceptual discrimination and also speed of response to stimuli. Surgeons are still the best agents for decision-making. Through our evolutionary past, we have developed an intrinsic ability for pattern recognition, decision-making and construction of plans, more flexible than any existing robotic hardware. With respect to our senses however, our performance generally is inferior for the detection and processing of incoming signals and, most importantly, for detecting small variations in the signals and the time to respond (47). In spite of being able to operate through two channels (thinking and acting simultaneously), humans are primarily single channel sequential recorders with a wide base and low sensitivity input/output capacity, with respect to their cost (48). The flexibility of the human response is limited and so operator-induced delay is of concern with respect to complex systems. Fortunately, however, there are few procedures in surgery where it is not possible to stop and apply an appropriate algorithm for analysis of the situation before proceeding.

Effects of fatigue: It is more critical to be able to recognise the need to apply such processes since, if the initial perception is absent or aberrant, then the appropriate path is unlikely to be followed. Humans are quite adept at storing multiple units of information, selecting, interpreting and acting upon this information. These functions are, however, decreased when fatigued or stressed and much of the input/output capacity is compromised. The natural predilection then is to revert to a more primitive mode, dealing with stimulating responses singly and sequentially, such that the competing signals need to wait for attention and a defined action. Sleep disturbance, a common problem in operating surgeons, can have major effects with malaise and a temporary reduction in intelligence quotient predominating, with consequential operating impairment (49). The fatigue, once established, is associated with a reduction in skill proficiency, psychological stress, resulting decrements of motivation and performance, often with slow, irregular and distorted performance. Experience suggests that decision-making requires cognitive information with the following components:

- Prior knowledge of the data source
- Memories of past or similar occurrences
- Simplification of heuristic rules (mental algorithms)

Inherent bias (50;51) suggests that heuristic rules and knowledge may be adapted and modified through training whilst memories are unlikely to be significantly effected.

3.5.4 User interface design considerations

Developing an original approach to surgical simulation involves accurately modeling the specific skills and requirements of surgeons. Those factors above outline the human side of the 'Human Computer Interface' (HCI) equation. The human performance aspects of this have far-ranging implications for the potential way that surgeons are screened in future for their training. The traditional teaching was that a surgeon should possess the *'Eyes of a hawk, the hands of a lady and the heart of a lion'*. This eloquent description by the fictional character Sir Lancelot Sprat in novels and on film (52) represents a romantic picture of the qualities required to make a surgeon, emphasizing high fidelity sensors and actuators.

The minimum entry criteria currently employed for surgeons, specify Basic Surgical Training skills as the only demonstration of physical and psychomotor ability required. Therefore the simulator design efforts have until recently been directed towards mastering these skills. *e.g.* VR MIST (17;43;53).

As detailed in Chapter 4, the planned evolution of simulation must take into consideration the educational application for which it is intended, closely linked to the surgeon's experience in clinical posts and their individual learning agreements with their mentors. Once familiar with an interface, the user wishes to maintain a degree of autonomy, modifying it and adapting it to their special needs and considerations. This process of customization should relate to the *user profile.* This should be integrated into both the evolving educational requirements and the development process for standards by which the systems are to be judged. Seymour (54) has recently developed evaluation methodologies for users, though this needs to be seen within the context of Higgins and Satava's (Figure 4.1) (20) work from the San Diego group. This is referred to as the validation model.

3.6 Evolution as a model for development of surgical simulators

The interfaces must meet the demands of the surgeon and the environment. This led to various approaches, based upon the technologies available at the time. The techniques being used to achieve pixilation through different means represent the diversity of evolutionary paths toward the same eventual 'goal' of an affordable high fidelity visual display system. For the concept of hybrid vision/graphic domains to evolve, convergent evolution of the different techniques would be required reflecting the 'tactic' of punctuated equilibrium (55), referred to as the iterative design. The process requires staged development with proof of concept at each stage.

3.6.1 The Process of Evolution- Iterative design

The system may be seen as an entity, constructed from its essential constituent components. This includes the use of various hardware and software components used in VOU1 and SIM1-2. It includes the technical design, and the use of *'customized off the shelf - COTS'* components where necessary.

The key point is that building a new simulation system requires surgeon involvement in the multidisciplinary team from the outset and so their limitations are seen as an integral part of the design. This included the initial work upon the development of shoulder arthroscopy databases established using the images of 'cadaveric' lamb and pig. Table 3.1 below outlines the main design phases and the associated trials.

Date	**Location**	**Project Title**	**Who**	**Trial**
1996	University of Exeter	Lamb World Model	Jason Cooper, Tim Bunker, Lindsey Ford	Alpha-testing only
1997	University of Exeter	Pig World Model	Jason Cooper, Tim Bunker, **Simon Grange**	Alpha-testing only
1998	University of Exeter	Human - Cadaver World Model	Jason Cooper, Tim Bunker, **Simon Grange**	1st prototype simulation (SIM1)
1999	University of Exeter	Human - Cadaver World Model	Jason Cooper, Tim Bunker, **Simon Grange**	2nd prototype simulation (SIM2)
2001	University of Bristol / PEOC	Flick Book - In vivo human	**Simon Grange,** Tim Bunker, Neill Campbell, Chris Setchell	Alpha-testing only
2002	University of Southampton	QuickSim1 constructed from the 2nd Prototype dataset	David De Roure, Nick Humphrey, Josh Burrill, **Simon Grange**	3rd prototype simulation (SIM3)

Table 3.1 Design phases of integrated simulation development

3.6.2 The iterative design process for simulation development

Three design phases are described in this thesis. It is inevitable that simulation systems will be strong in some attributes and weak in others. With video based modelling developed in the *1st design phase,* the problem of limitations of interactivity within the environment forced redesign in the 2nd phase. The eclectic approach for the evolution of a development strategy encompasses the use of computer science methodology and ongoing performance evaluation of surgeons. Comparison with other such systems needs a benchmarking process.

By the 3rd prototype (SIM3), integration of the evolved simulation system within a training package (Table 3.2), professionally supervised by peer group experts, was developed with Java and hypermedia development tools. This was prepared within the framework of a tutorial that was trialled by appropriate groups of participants (trainee surgeons). The results were published on-line as part of the evolving dynamic review journal (56). The intention is to use this training system as an adjunct to the standard training courses in specialist areas of surgery. The later VOU2 prototype provides a benchmark for standards setting in this field. Ota (57) proposed the potential value of such simulators in 1995.

Date	**Functionality**	**Scope**	**Validation state**	**Integration**
SIM1 1998	Video based simulation with tracking (SGI)	Human Shoulder	Small group evaluation (Pilot study)	1st prototype simulation (VOU1 & SIM1)
SIM2 1999	Video based simulation mouse driven (PC)	Human Shoulder	Small group evaluation (Pilot study)	2nd prototype simulation (VOU1 & SIM2)
SIM3 2002	Embedded simulation mouse driven (PC)	VOEU (Hip, Knee, Shoulder)	Small group evaluation (Main study)	3rd prototype simulation (VOU3 & SIM3)

Table 3.2 Functionality & scope of prototypes

3.7 Virtual environment model design

There are drawbacks with the viable alternatives to the development of simulations at present, generated from digital virtual environments (VEs). There is an ethical obligation to develop tools as efficiently as possible. This requires triaging of the process to optimise the limited resources, prioritised through the series of experiments and iterative design cycles. This includes software components, hardware platforms, networking and consideration of the formalised training schedules, and also advances in surgical engineering, bound by the limitations of healthcare system training resources.

Task Analysis: Task analysis of the surgical psychomotor sub-skills was achieved in conjunction with the Helmholz Institute in Aachen for the 2nd EVW prototype but little change was made to the system from the original concepts developed by the author and Tim Bunker for the first prototype. This led to greater consideration of User Needs Analysis / Usability Analysis in the design of prototype 3 and was based upon work with Dr Gary Wills at the IAM group in Southampton. Upon each occasion, what appears to be a relatively clear-cut development timetable has been significantly altered as a result of seeking funding or availability to resources, particularly clinical facilities.

The aim of the design of prototype 1 was to generate an adequate illusion of a moving image. A sequence of correct digital video images was selected by the computer and presented in relation to the observer. As the surgeon moves through the environment the sense of movement is generated by the continuing sequence of video images shown on the monitor, much as if one were looking at the monitor of a normal arthroscopy. This system worked well, aiding in creating an adequate illusion because the procedures are immersive, in that the operator is normally working from a TV monitor.

3.7.1 Computing issues in development phases

Experts from the many different centres of expertise should always define computer resource requirements. The initial influence of ATM directed by support from K-Net Ltd[11], who sponsored early Initial Prototype work, has been overshadowed by the failure of UK healthcare institutions to take up the technology (only 2% of hospitals) and so the funding shifted emphasis.

Hardware costs have fortunately decreased partly in line with Moore's Law's expansion of technological performance (58), making this less of an issue as the projects develop. The task analysis, architecture, and internal computational requirements have been discussed in the multidisciplinary meetings with the members of the various teams mentioned above in Chapter 2. Software production has been in-house, except where stated.

[11] K Net Ltd. http://www.k-net.co.uk MD Mr Henrik Kjellin

3.7.2 Educational issues in development phases

Educational Infrastructure Requirements are based around the Wessex Region for Specialist Registrar training in Orthopaedics and the policies of the Royal College of Surgeons of England. Authorisation has been by tacit support for the research work over the last 7 years with input from trainees in the Wessex and Far South West regions. The task analysis has been expanded as part of earlier work in prototype 2 – worked out with Lucile Vadcard (UJF - Grenoble) and Prof. Grainne Conole (University of Southampton) as part of the VOU2 pedagogy development as described in Chapter 5. Development of the approval strategy for simulations is still in progress and the AIPES classification system (59;60) proposed here is an integral part of this for VOU2. This sets the baseline validation criteria.

Expert defined clinical content was used for the development of the tutorial section of the prototype 1 and in prototype 2. It was based upon the earlier book by Mr Tim Bunker (35). Tutorial material developed for the third prototype which is stored in the library section of the VOEU shoulder module, is based on a later book also by Mr Tim Bunker (61).

3.8 Validation issues in the development phases

Decision-making regarding the simulators' effectiveness as educational tools is going to depend upon the continuing evolution of the 'Auto testing' of simulator performance through continuous data collection on user activity and opinion. Ultimately this may lead to a comparison between simulators versus traditional testing methods. The approach in the trials (for prototype 3) was to use a questionnaire that was available as an active server page run on the *Microsoft.net*™ server and the collected data is then analysed using the DRJ. The results of this are discussed in Chapter 8.

3.8.1 The simulation performance evaluation paradox

This simulation system will ultimately offer the ability to obtain objective quantified data as to the technical performance of an individual surgeon. Therein lies a major challenge. It is currently impossible to evaluate sensibly surgical simulators since there is a self-referential paradox that underlies the problem.

One cannot test a simulator unless one has a known population performance

One cannot test a population unless one has a known simulator performance

Expanded, one cannot assess a simulator unless one has a performance standard for the population to be tested. However, one cannot reliably test an individual unless one knows the reliability of the system (simulator) used to test them as compared to an average population, since this is necessary to fairly discriminate individual performance.

3.8.2 Training and Certification of the Simulator

To an extent, this depends upon documentation. In order to support the development of Training Courses Integration, the author has prepared the following documents:

- **ICT Skills for healthcare professionals** book - Appendix 2.1, aimed at providing a bridge for newcomers to the digital domain.
- **Curriculum integration proposal.** This was prepared as part of the second prototype and has since been surpassed by other VOEU educational design documents, but represents effort by the author to address the curriculum issues in detail.
- **Publication of results as a dynamic review.** This is an ongoing process as described above and the trials relate to the third prototype.
- **Evaluation and performance specifications for a surgical simulator -** Each version of the simulation prototype has different purposes and thus different design requirements.

Three issues arise:

1. **Visual acuity –** The database was adequate to meet the human visual performance limitations & demands outlined in Section 3.5.1. The system achieved a latency of less than 50 ms^{-1}.
2. **Tracking systems –** cost is restrictive and access limited in United Kingdom Hospitals presently imposing critical factors upon uptake and evaluation.
3. **4 *dof* transition to 2 *dof*** databases for the third prototype. This decision was made to complete a working prototype within the project limitations.

3.9 Subcomponent Performance – Prototype 1

The development of virtual environments (VEs) constructed from video material was outlined in 1995 (62;63). Prototype one played back good quality visual images to the observer in real time. An example of this system in operation is described elsewhere in more detail (36;64). Setting a practical threshold of around 10 hours for each 'virtual world' for prototype one, the virtual world was made up of video capture of around 5 Gb of video frames filmed with a frame rate of 10 fps. This fits with a 1mm x 1mm x 1mm grid (x, y and z planes) and a 5-degree rotation (plane of 360 degrees) allowing 5 – 10 cc of joint space to be recorded. This should allow for a shoulder joint to be adequately surveyed, although this may cause problems with the complete filming of a knee joint, in particular if there is significant fluid distension of the suprapatellar pouch increasing the volume.

These parameters are likely to be reset again as advances in data storage for video recording and access, becoming both cheaper and faster. The ‘Holy Grail’ of photo rendering, matching the visual acuity of the operator, is unlikely to be achieved until at least 2007 based upon the predictive strength of Moore’s' Law. Progression toward larger databases, which will expand to include the common pathologies recognised as essential for the training of individual surgeons is one option. This was explored for the third prototype, which demonstrated how to build an atlas of pathologies using intraoperative video. Another option would be to explore parallel image transformation algorithms such as exploiting salient stills technology, and 'morphing' from key frames.

These decisions result from the extra complexity that this simulation imposes. This is too great a technical and cost burden upon the end users system to make it a seriously viable option. Although compromising the effort expended, the decision to proceed with streamed networked video, running over hospital networks, supporting access to the Internet was a pragmatic approach to the trials and further iterations of design.

The *Exeter Virtual Worlds (EVW)* System video world infrastructure used various hardware and software components (detailed below) in these early experiments (1996 - 1999). This highlights the relevant advances made by this early work, resulting in the development of the prototype human cadaver modelling system. During the early experiments, the Exeter Virtual Worlds Team (36;65) included colleagues such as Mr Jason Cooper who, with the author initiated the first biological experiments.

3.10 Prototype one equipment engineering

Considering the process of database material generation, the laboratory set-up that was necessary for the virtual environment model design consisted of two core fields of technology, Biology and electrical/mechanical engineering. *Figure 3.3* demonstrates the equipment. An environment that will not adversely affect the tracking equipment is essential. The following notes detail the experimental environment issues.

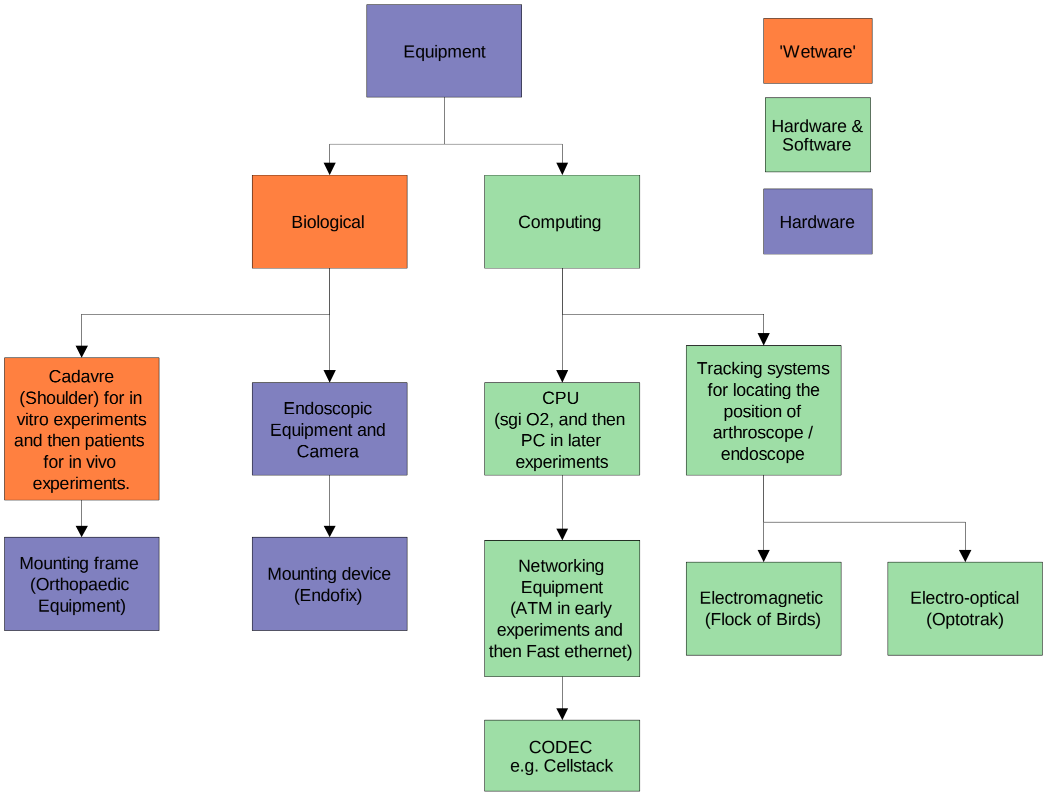


Figure 3.3 Recording system laboratory equipment used in the preparation of shoulder virtual environments

The following factors affect the quality of recording:

- **Cadaver.** Critical factors are the cadaver's freshness, its completeness, and whether there is a fluid level in the joint, as some surgeons prefer a 'dry scope' at least initially, and the entry portal is crucial. This effects the positioning of the scope as outlined in the VOEU on-line handbook[12]. Finally, all biohazard limitations, especially with respect to handling human tissue, must be respected.
- **Camera.** The camera mounting affects the accuracy of the mounting base co-ordinates, and any future experimenters should consider the issue of whether to use manual or robotic control of the image collection platform. For these experiments, a manual base was used. Tolerances for the acceptable image granularity matrix must be defined (1*mm* depth measurements for the image z-plane is the coarsest range in the cadaveric experiments), and also adequate degrees of freedom need to be confirmed.

[12] Detailed description of the posterior portal anatomy is available at the VOEU site at: http://voeu.ecs.soton.ac.uk/voeu/library/doc.aspx?docid=68

This includes adequate range of arc, which was restricted by the isocentric mounting device. It is important to check for consistent errors, biocompatibility of the mounting, and ensure the camera quality. A standard analogue arthroscope *Olympus* video camera was used in the experiments to obtain the initial images that were then stored using (JPEG lossy – q50) compression on the hard disk.

- **Filming Process.** This can employ analogue or digital 1 or 3 chip signal-to-noise ratio filters for recording of the video stream data using a separate hard drive, with 20 Gb capacity as the storage medium. The bandwidth to support the flux of the data stream is important. With local filming this tends to be easily achievable usually requiring less than 10 Mbs^{-1}.
- **Quality of reproduction.** This relies upon the degree and type of *(lossy or lossless)* compression. The data storage medium, its durability and the replay system need to be taken into consideration. The reliability of reproduction will depend upon the realism of output devices, markers of reality and the tracker accuracy.

The author was heavily involved at this stage of the Exeter Virtual Worlds System design, development and testing, becoming more directive in later stages of design. Iterative design cycles were similar in nature, once the methodology was defined. The process of building a virtual world requires the computing and biological hardware, and the computer system architecture. Each is detailed below.

Figure 3.4 Exeter Virtual Worlds Simulation System is specifically designed to teach pattern recognition skills in surgery.

3.10.1 The Computing System Architecture

The system used five technologies:

i. Digital video
ii. ATM networking
iii. Multimedia databases
iv. Position tracking
v. Client/server computing

More information regarding the theory of simulation using this system is available in other published works on the topic (36;66). An analog stream of video information can be digitised for ease of transmission and manipulation by digital computer networks. Initial experiments used a dedicated networked device called a video coder/decoder or 'codec' to perform the digitisation and reconstruction of the video images. Each frame of the stream of video can be referenced and manipulated separately. Digital video has many uses, *e.g.* distribution of live images of operations for training purposes or archiving complex real time imaging such as cardiac catheterisation.

3.10.2 ATM networks

The network requirement was for a real-time and guaranteed bandwidth transport between the database of video images and the video codec. The asynchronous transfer mode (ATM) network provides these characteristics and the bandwidth (or speed) of the network can be scaled as needs arise. ATM networks were originally chosen as they are very robust compared to legacy equipment and can run all the usual legacy systems (*e.g.* Novell, NT and the Internet Protocol: supporting academic and administrative infrastructure) in parallel with video, audio and control information. In addition to the video data sent from the multimedia database to the video codec, control (narrow bandwidth) information is sent from the operator's positional electronics (the Bird) to the database, indicating which frame to retrieve next. This '16 port' Switch system offers guaranteed bandwidth information transfer at a speed of 155 MBit both ways, receiving inputs from Fibre or UTP cables, allowing for good integration with IP. ATM programming consists of relatively simple streams, like TCP. Detailed description is beyond the scope of this thesis.

3.10.3 Recording the image database

Video images were stored and retrieved from their associated co-ordinate positions. The video images are stored as sets of network messages, directly relayed to the video codec to produce the images. The database associates the video frames with the co-ordinates from where they were recorded. During playback, the co-ordinates of the camera are used as a key into this database with inexact matches being arbitrated by Euclidean mathematics, that is, the 'nearest' frame is chosen. The database allows for the random access of the images.

To collect images from the joint it is necessary to have a fixed fulcrum around which the arthroscope rotates to allow for translation corrections when the image is mapped onto a database by virtue of its coordinates.

The use of a fixed axis ensures the images are collected radially from a fixed point. To space the z co-ordinates equally from the centre of the recording outwards, plastic spacers of increasing size (1 mm increments) were placed as 'shims' to ensure uniformity of image capture. This allows for the operator to film a 'grid' which for each given y co-ordinate has an x co-ordinate recorded as the operator moves the arthroscope across the field. Adding spacers increases the depth of the model so that the layers (of *z* co-ordinates) are built up much like the layers of an onion.

This correlates with the geometry of a computer's hard disk and makes retrieving of adjacent frames a matter of moving to adjacent tracks on the disk. While enhancing the speed of playback 1.7 fold (as opposed to randomly distributed data), the search for the next best frame to display, can also be localised based upon the speed of movement of the 'virtual' arthroscope. To search all frames would be too slow, thus the algorithm is generated to 'look' for the frame that is likely to be nearest should the current path and speed of the 'virtual' arthroscope be maintained.

3.10.4 Tracking systems - the flock of birds

The Bird™ is a 6*dof* sensor (locator) device for tracking the arthroscope. It is a serial device that can allow for 6 degrees of freedom: Euclidean 3D, and also roll, pitch and yaw. Only the first three were used in the final teaching models. A 4+D system ultimately will be required in order to accommodate the rotational component of the arthroscopic technique.

The Bird™ was used to identify an object's location in 3D space (normally described as *x*, *y*, *z* Cartesian co-ordinates) and its angle rotation against these planes (x′, *y*′, *z*′). It does so through transmitting and receiving a signal using low energy electromagnetic radiation. Its effective range (radius) is approximately 1 metre. The angle of rotation against the x plane is called *pitch*, the rotation against the y plane *yaw*, and the rotation against the z plane *roll*. This acts as a generator of co-ordinates, which acts as a reference to each stored frame.

Six degrees of freedom are required to completely specify the location (3D) and orientation (roll, pitch, and yaw) of a surgical implement held in Euclidean space. However, when this is a virtual arthroscope entering or rotating via a fixed portal, 4 (four) degrees of freedom are adequate to specify its positioning as the other (z co-ordinate) is fixed.

3.10.5 SGI Indy/Indigo silicon graphics workstation

The computer initially used in experiments was equipped with 96 Mbytes RAM which could be used to cache recent frames. Only 10 Mbytes of RAM were used with striped 1O Gb hard disks. With a 10-kbyte-frame size this allowed for 1 million frames on disk. Using this system, there was *no* use of the computer's graphics hardware. As systems developed it was possible to use IBM-based PCs in the second prototype and a dumb client in prototype 3. By this stage, physical storage was no longer an issue.

3.11 1st Prototype biological hardware

Earlier experiments to develop the technique relied upon pork and lamb specimens. This was economical and ethical, as specimens can be obtained from a butcher and the methodology for filming could be optimised without jeopardizing very expensive cadaveric material. This does however pose a major disadvantage. The anatomy of the joint is not entirely representative of a human and therefore cannot be used for training.

For the cadaveric experiment a practical sized database was adopted. The biological limits will be critical in the future especially for intraoperative recording. For the experiments, unfixed cadaveric material was used, mounted using an orthopaedic fixation device, so that access to the capsule (covering the joint) over the portal (entry hole for instruments including the camera) was readily available. This acts as the axis about which the arthroscope moves and poses a potential source of error for the motion tracker co-ordinates, which are reciprocated from the other end of the arthroscope. The specimens can be positioned to allow some distraction of the joint space[13], otherwise the architecture is unaltered, and irrigation (normally used in the surgery) was not used in the experiments developing prototype one. The following biological hardware inventory was required for the experiments:

- **The arthroscope.** For playback it is possible to use a pseudo-arthroscope, which can be simply a rod or pointer, which has the sensory input of a tracking device attached to it. Ensuring that the weight of this structure is similar to that normally used by the surgeon considerably enhances the sense of presence. In fact in the experiments an actual arthroscope was used.
- **The monitor.** The colour monitor is used to observe the arthroscope camera output. High quality, HDTV output is normally generated. This image is full stream, full screen, and the operator is normally unaware of delay in image processing. This reflects the quality of the analogue camera systems used, because image processing does occur,

[13] The joint surfaces are normally in contact. When operating, fluid is injected into the joint to move the surfaces apart so that a clearer view can be obtained and so as to minimise the risk of damaging the surfaces. In the case of the experimentation this was achieved mechanically rather than by fluid distension of the joint cavity.

for example to modify attributes of the image such as colour balance[14]. Later experiments employed single chip digital cameras.

- **Location tracking device.** Attachment of the tracker device to the end of the arthroscope adequately mimics the camera lead. The virtual worlds were filmed with both the Flock of Birds™ and the Optotrak™ devices attached to allow for comparison between the two systems. Only the Bird™ was used for playback.
- **Reference frame.** A purpose-made secure mounting system was used to hold the arthroscope during filming, allowing 4 degrees of freedom of movement and providing a fixed location for the fulcrum to minimise error in co-ordinate recording.

3.12 1st Prototype methodology for the virtual environment building process

The process of video-based virtual world capture has three distinct phases:

- **<u>Step 1</u>: Recording a world.** Firstly fixed points, such as the portal of entry into the joint, are recorded and then all possible views from that portal, guided by an 'in house' Capture Visualisation 3D mapping program, were recorded.
- **<u>Step 2:</u> Computer analysis of video database.** Over-sampling during recording could lead to more than one frame being available for selection during playback. The software allows for interpolating co-ordinates and selection of the best frame.
- **<u>Step 3:</u> Playback of the virtual world.** Playback involves using the Bird transmitter, which is moved whilst it is attached to the arthroscopic representation. Repeating the sequence of VR input, frame decision and frame output generates the ‘moving’ image.

3.12.1 Search strategy developed for frames

Grid search errors could be introduced by using a search based on the geometric partitioning of the search space, since the track of the user will not necessarily be similar to that of the original ‘frame recording’. In effect the frame selected is the computer algorithm’s ‘best guess’. However, the impact of this is not always apparent to the user. The search locality is based upon the speed of the operator’s movement. To minimise the time required for this in real time play back, pre-computation of all co-ordinate choices would have been an option.

---

[14] This can be demonstrated by rotating the camera rapidly to emphasise the individual frames, which takes so short a period as to induce a lag of less than 0.1 of the second. This would eliminate its perception in the normal operating environment, where such sudden (ballistic) movements are seldom intentional.

3.12.2 *Discussion of initial prototype*

The methodology of the task analysis for the 1st prototype simulation trial was based upon the practical limits of what could be achieved with equipment then available, and also the expert analysis of tasks (35;61). The results were inconclusive with a small number of trainees (N=16) initially tested in the pilot study, but this did however provide a foundation for developing the next prototype and for conversion to a PC platform, and also *in vivo* recording. It was anticipated a larger number of trial candidates would be required, as has been described in similar simulation trials (67). This only provides information upon sub skill analysis.

The impact of the early experiments during the prototype was to minimise the effect of the granularity of the database (*i.e.* resolution) based upon this key perceptive issue. This was one of the factors affecting the level of detail in the simulation. The resultant restriction of the numbers of frames necessary is in Appendix 3.1. The ultimate goal would be a dataset with a granularity finer than the limit of human visual resolving power with normal visual acuity; hence the need to define the visual threshold and passing an A-B comparison test. This is dependent upon the distance of the observer from the virtual object, and other factors such as light levels and contrast of colour. To meet the storage demands for a complete dataset of a specified resolution a calculation of the database size is described by the formula (68):

$$\textbf{Required Space} = \textbf{Frame Size} / (\textbf{Frame Compression} \times \textbf{Number of Frames})$$

Equation 3.1 System Storage Demands

The time available for data acquisition is dependent upon the use of fresh biological material which imposes potential limits due to the biological decay rate in environment (see Page 38). This defines the absolute time. The potential development of a biohazard was used rather than time taken for the biological image to decay. This can be described by the formula:

$$\textbf{Number of Frames (f)} = \textbf{Frame rate (fs}^{-1}\textbf{)} \times \textbf{Absolute time available (s)}$$

Equation 3.2 Limitation upon filming time

The tissue quality degrades exponentially after a certain time due to the biological process known as necrobiosis. This results from bacterial agents destroying the tissue. They increase their population size after an initial 'lag time' of about 6 hours and influence the rate of decay. **Biological Decay Rate** (%/hr) depends upon other factors such as ambient temperature, which affects bacterial growth. After about 6 hours bacteria display exponential growth producing destructive chemicals leading to decay in quality and increase in biohazard for non-fixed specimens. Lowering ambient temperature can prolong the time available since this slows the rate of bacterial reproduction and hence the size of the population.

3.12.3 Disc space and leverage

This was an issue in the early experiments with databases of around 1 Gigabyte, though this is now trivial and multiple databases of this size can be stored. This is only relevant when considering its distribution mechanism, where 'store and forward' or appropriate compression strategies for streaming are implicated. There are many ways of setting up the recording device at different distances from the fulcrum, so that one can increase the resolution of the images or gain a greater range of movement within the joint.

3.12.4 Frame size

Though other combinations of frame size and compression were tried, the hardware was found to tolerate the 250Kb image compressed 25:1 without error, resulting in a compressed frame size of 10Kb and allowing real-time playback without difficulty. This provided adequate visual resolution at a refresh rate of 25 frames per second, which is the best performance available upon this hardware set-up. Subjectively users actually found the jpeg compression technique enhanced the images by accentuating boundaries. Storage required can be described by the formula:

Storage capacity required (Mb) = Frame Size (10Kb) x Resolution (mm^{-3}) x Volume (cc)

Equation 3.3 Frame Density Formula

Estimations of the approximate database sizes are given in the Appendix 3.2. This is demonstrated in Table 3.2 that combines the need for frame number per volume *(cc).* This is the image density. The estimated volume is 10 cc for the shoulder joint, 15 cc for the subacromial space that is above the shoulder joint and 25 cc for the Knee joint[15]. This is easily measured and the standard compression ratio of 25:1 and 10 Kb frames allows for a visual granularity of various ranges from 2 mm to 0.33 mm between each frame with different rotation values, expressed in Gb of hard disc space required for storage. Seen in the context of the capacity of the joints to be modelled, it is necessary to explore the relationship between recording rate and hard disc space since this will dictate the joint capacity that can be mapped, as referred to in the table below. In fact the system optimally supported recording at 7 frames per second, but in order to assess which frame density would be practical to record, a table of calculations (Table 3.3) was made.

Frame rate of recording / Storage planned	**1 Gb**	**5Gb**	**10Gb**	**20Gb**
5 fps	4.2 hours	*21 hours*	*56 hours*	*102 hours*
7 fps	3 hours	*15 hours*	*30 hours*	*60 hours*
10 fps	2.1 hours	*10.5 hours*	*28 hours*	*56 hours*
15 fps	1.58 hours	7.87 hours	*21 hours*	*42 hours*
20 fps	1.05 hours	5.25 hours	*14 hours*	*28 hours*

Table 3.3 Recording duration related to planned storage. Figures in italics are not feasible.

3.13 Practical set-up of the simulator development laboratory

The equipment and experimental techniques used to develop and record the image database necessitated that consideration be given to the risks of contamination and restrictions upon the time available for the database preparation. This raised key issues for future development of models using this technique. It was envisaged as a stepping-stone to building a digital atlas of clinical cases and therefore *In Vitro* perfection of technique was necessary before *In Vivo* data collection was commenced for databases for prototypes 2 & 3. At the heart of the development of the recording phase was the need to collect spatial co-ordinate data along with the visual data. For the purposes of these experiments, equipment was organised for three specific applications:

[15] This is easily measured. The joint capsule is relatively inelastic so once it is filled to capacity with fluid, the resistance to stretch increases. Therefore, the pressure required expanding the space increases dramatically (Boyle's Law). In view of the fact that the pressure of the irrigation fluid is set to generate a positive pressure inside the joint, forcing fluid out, the capacity of the joint is relatively easy to measure.

- Image collection in a laboratory
- Image collection in an operating environment
- Replay of the database in a training environment.

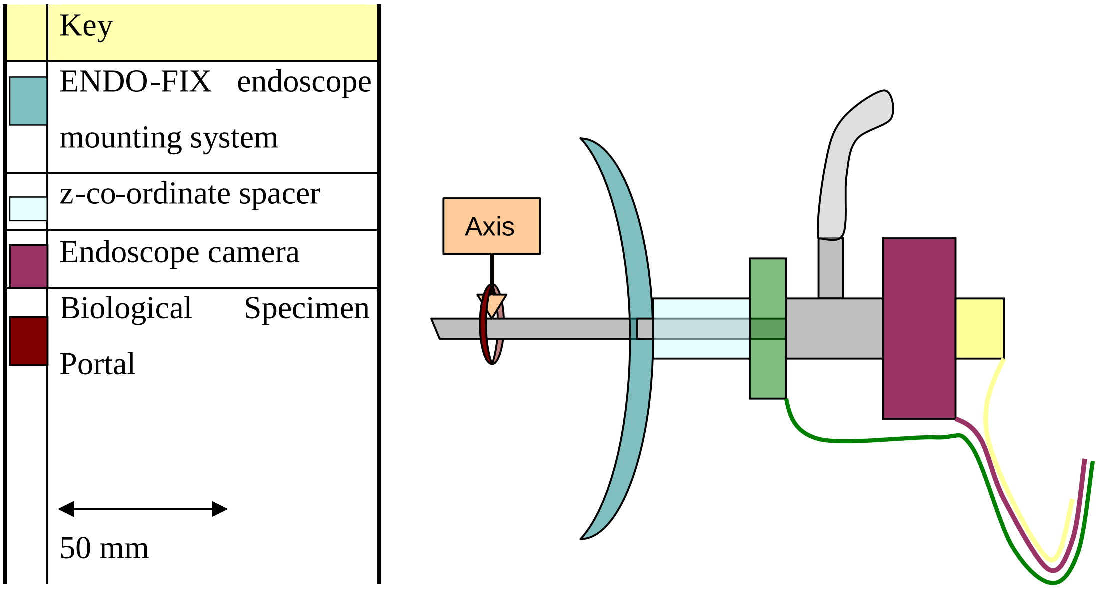


Figure 3.5 Video camera and arthroscope making up the endoscopic recording system

3.13.1 Endoscope holding device - ENDO-FIX

Originally designed to achieve good fixation so as to deliver a steady picture, using an endoscopic camera (Figure 3.5) in an operating theatre environment, the endoscope-holding device ENDO-FIX (Figure 3.6) was developed at the Forschungszentrum Karlsruhe *(Patent DE 4413918).* It is the key component system in this rig.

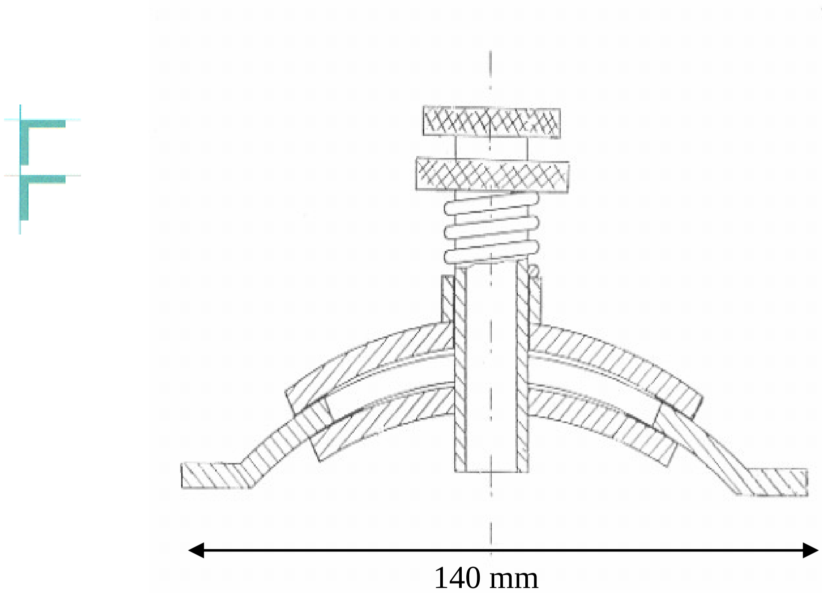


Figure 3.6 Endoscope holding device - ENDO-FIX

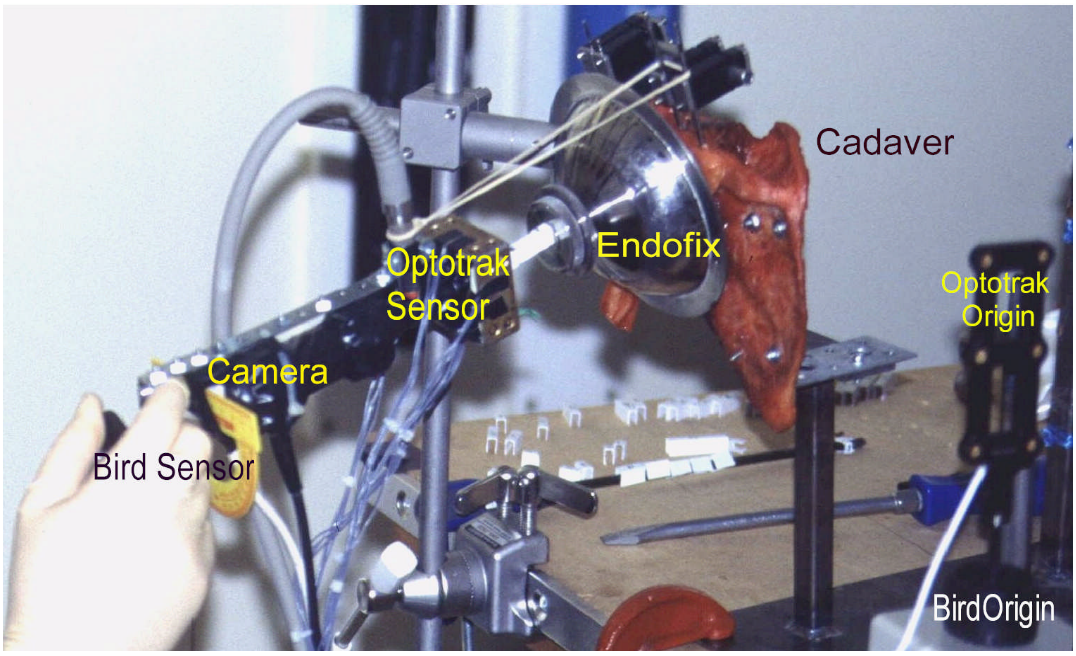


Figure 3.7 Experimental rig with cadaveric shoulder specimen in situ

3.13.2 Comparative measuring systems for spatial tracking

Two comparative spatial co-ordinate referencing systems were used to identify the position corresponding to each individual frame of the recorded video database *(Figure 3.7)*. In this case the Electro-optical (Optotrak™) spatial reference system is compared with an Electro-magnetic system (Flock of Birds™). The indication for using the Optotrak™ system was to obtain an accuracy of registration in the order of the 1 millimetre level – which is the same as the granularity of the database in the *x, y, z,* co-ordinates. This ensured accuracy on a par with the clinical procedure, allowing for leverage around a 'fulcrum' (the axis of where the arthroscope penetrates the joint). This also serves as a preparatory environment for the assessment of equipment that may be used in the operating theatre environment, to accurately record the spatial co-ordinates of pathological images.

Electro-optical referencing systems can offer greater accuracy than comparable electromagnetic systems. Due to the nature of the energy transfer used to record positions the system is well suited to the more complex operating theatre environments where potential electromagnetic interference with other critical systems can be avoided. Such tools are already used within the operating environment for *image guided orthopaedic surgery (IGOS)* (69-71). The high capital cost may however prohibit the widespread use of such equipment for training (around £25,000 - € 40,000 for a complete hardware and software system in 2006). Thus by developing mathematical correlation between the two systems it is possible to consider their specific use in different environments. There are specific issues of calibration and registration discussed below along with analysis of the mathematical correlation of the two systems (72).

3.13.3 The Optotrak™ & Flock of Birds™ Considerations

The Optotrak™ requires a stable platform between 1.2 m and 1.8 m above a floor that will maintain the position of the Optotrak™ within an accuracy of +/- 0.3 mm and support its 100 kg mass. It should be situated not more between 2 m and 4 m from the specimen (*i.e.* the markers attached to the arthroscope whose position it records).

The Flock of Birds™ (Ascension Technologies, USA) offers control of up to 12 points in space for co-ordinate plotting with 6 Degrees of freedom *(dof).* Only one point in space, the position of the arthroscope, is recorded in these experiments. This provides a relative coordinate matrix for the image frame positions. It is a serial device. Euclidean 3D and roll (α) are recorded to provide a position for the arthroscope. Its effective range is approximately one-metre radius. However, it can generate co-ordinates of 2 mm accuracy within this range, based upon the position of the camera end of the arthroscope (Figure 3.7). The Flock of Birds™ requires that the base unit is situated within 1m of the sender unit (attached to the arthroscope), and for practical purposes the movement of the "Bird" remains within one hemisphere to minimise mathematical readjustment of the co-ordinates during playback.

3.13.4 Mathematical analysis of location tracking systems

The two tracking systems operate in different ways. Analysis of relative and absolute Euclidean geometry has been carried out in association with Mr Jason Cooper. In order to correlate the position of two locators (markers) there needs to be a mapping function that ensures that the origins for both locators are coincident. Both locators have the same orientation of axis, their roll, pitch and yaw values are the same. When *the two systems are axis-parallel and the order of rotations is the same,* both locators use the same units of measurement.

If any or all of these conditions are not initially met, the co-ordinates of either locator can undergo ‘pre-processed’ transformations to meet the corresponding condition(s).

In the specific case of 3 dof (x, y and z) vectors where multiple translations are required, their order does not affect the result since they are *linear*. If an object is rotated and then moved, its new position would be identical if it were firstly moved and then rotated. The process can be described in the following stages:

1. Map the Optotrak™ sensor marker co-ordinates to the sensor origin co-ordinates. The algorithm proposed refers to the two sets of co-ordinates for the bird origin and the bird tracker.
2. Inverse/reverse transformation of the Optotrak™ marker to the Optotrak™ origin (point of reference). This is an Inverse Transformation. The actual data frame is mapped to the origin in inverse order by the following steps:
    - 2.1. Translate in negative x,y,z-directions
    - 2.2. Rotate by negative angles and in reverse order (rot z, rot y, rot x in contrast to forward transformation) for the Optotrak™ Marker co-ordinates to move them to the Optotrak™ Reference co-ordinates.
    - 2.3. Map the Bird™ sensor marker co-ordinates to the sensor origin co-ordinates. Inverse transformation of the Bird™ marker to the Bird™ origin (point of reference).
    - 2.4. Correct the offset between the Birds’™ marker and the Optotrak™ marker system, since the two markers cannot occupy the same position on the arthroscope simultaneously. A simple mapping function can then be used for playback with the Bird™. This process is demonstrated in Figure 3.8.

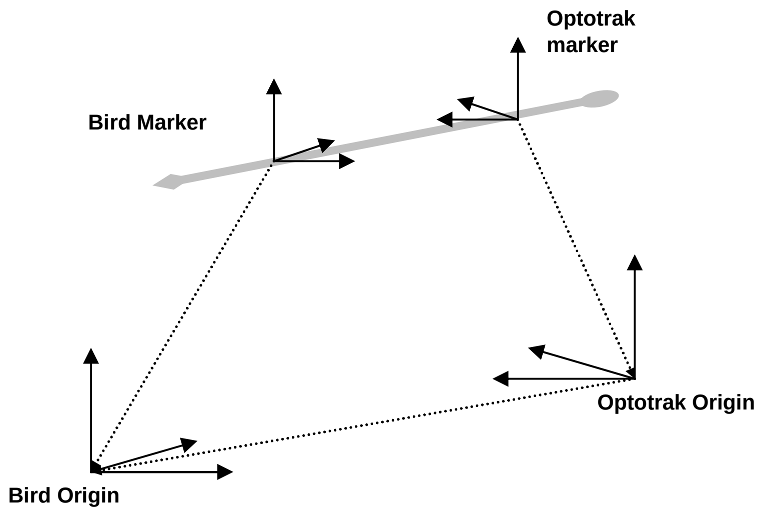


Figure 3.8 Local and global matrix transformations between co-ordinate sets. To simplify the diagram, the co-ordinates do not show rotation. Since the arthroscope is inclined by 30 degrees, this will rotate the local co-ordinate systems of Bird™ Locator and Optotrak™ marker systems against the original (reference) systems of Bird™ Optotrak™.

3.13.5 The simple mapping function

Formally, the above will be expressed in the following way:

When two objects O and O' are connected by a beam of fixed length (l) and the 6 *dof* co-ordinate location of one object is known, that of the other can be deduced, allowing a one-to-one mapping to be established between two locators rigidly connected to the same arthroscope. The simple mapping function can only be used after the pre-processing has been performed.

Let the position of the first locator **O** be **p** = (x, y, z, a, b, c) and the second locator **O'** be **p'** = (x', y', z', a', b', c'). The two locators are a distance 'l' apart which can be expressed as a displacement vector **d** = (x_l, y_l, z_l, 0, 0, 0) (Figure 3.8). Now, **p'** = **p** + **d**, so from the above: **p'** = (x +x_l, y + y_l, z+z_l, a, b, c), where x_l, y_l and z_l are the only unknowns. From Figure 3.9, these can be worked out by geometrical arguments to be:

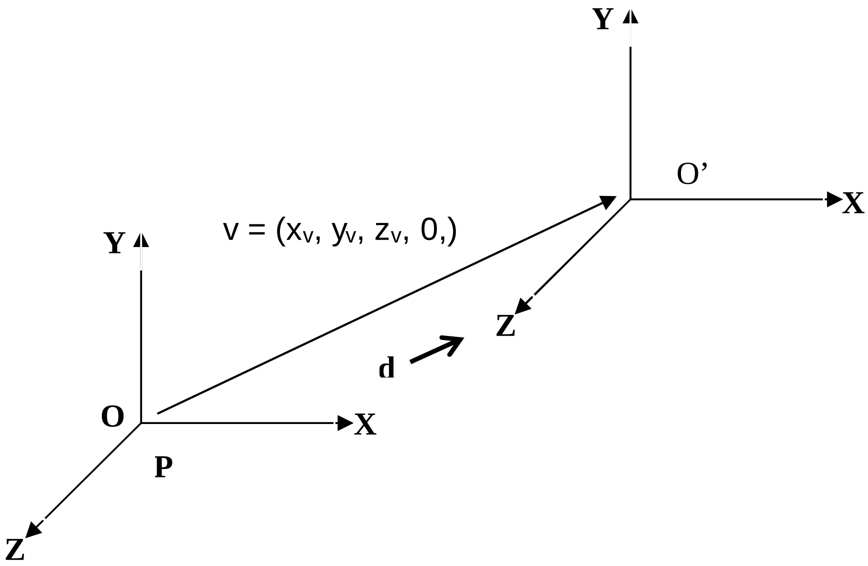


Figure 3.9 Translation between two relative co-ordinates

$$x_1 = l \sin(b) \cos(c),\ y_1 = l \sin(c) \text{ and } z_1 = l \cos(b) \cos(c)$$

Equation 3.4 Simple mapping function

3.13.6 Alternative transformations of the recording matrix

There are two alternative ways in which the co-ordinates can be obtained from each locator by transformations (required for *calibrating*).

- **Origins:** The first type of transformation involves ensuring that the origins of each locator are co-incident. In practice, this type of transformation will always need to be performed because the origins are physical devices, which cannot be located in the same physical position. In general, an offset vector **v** = (x_v, y_v, z_v, 0, 0, 0) can be subtracted from one of the locator's co-ordinates to give co-ordinates expressed from the other locator's origin, (Figure 3.9).
- **Orientation:** The second type of transformation ensures that the orientation of both locators and origins are identical. As with the first type of pre-processing, an offset rotation vector **r** = (0, 0, 0, a_r, b_r, c_r) can be added to one of the locators to orientate it in the same way as that of the other.

3.13.7 Comparing the two tracking systems

The need to move the origin and then record the origin suggests a need to monitor the origin, rather than simply take point observations. If this is the case, it may be better to use a separate Bird and calculate the relative distortion of the EM field. This is however only a consequence of the limitations of the recording apparatus. In an ideal situation one would be able to record and play back material without incurring registration errors, using accurate and cost-effective hardware.

Electro–optical equipment can be used in the operating theatre environment. This is considered a critical environment where high-energy electromagnetic (EM) field generation may cause unwanted interference within the environment affecting other critical systems such as monitoring, although there are clear range considerations since field strength decreases by a factor of $1/d^2$ that may still allow EM radiation within certain areas of the critical environment. The Bird™ should therefore remain at least 2 metres from any electrosensitive equipment *e.g.* anaesthetic monitoring.

EM devices are affected by the ferromagnetic characteristics of the environment. The Bird™ will record an offset of co-ordinates due to the use of a recording frame, which has a high ferromagnetic content. The advantage of using such equipment for play back in a similar environment means that the errors in the field matrix are cancelled out. Unfortunately, this is not the case when the playback set-up is a different '*ferromagnetic environment*'. Newer 'laser bird' tracking systems may be employed. This type of resource was not available to the project. To facilitate the possible integration of such systems it would be necessary to first ensure that it is possible to register the recording and playback systems relative to each other. In view of the greater accuracy of the Optotrak™, which is reported to be accurate to within 0.33 *mm* in the x, y and z, planes, it would be logical to use this as the base line reference system. It does of course bring into question the underlying accuracy of the Optotrak™ since this implies that it has become the gold standard, acting as the benchmark, and thus any aberration between its reference matrix and the real world will not be detected within any further studies. There are two reasons for adopting this approach:

- The author is unaware of any systems currently either commercially available or in research laboratories (see Figure 3.10 for laboratory environment) which offer greater accuracy than the Optotrak™, and in this experiment it is satisfactory since the error is less than that induced by the amplitude of physiological oscillation of the manual operator.
- The precision of the equipment exceeded the original specification of the video world database matrix granularity, set nominally at 1 mm, to ensure that the database

could be stored on hard disc. Operatively, the leverage of using an instrument via a portal is taken into consideration.

3.13.8 The relative merits of tracking systems

The Optotrak™ system allows the accurate registration of the frame recording position, which can then be recalled using any system providing suitable accuracy for the operator, obtaining access to the part of the database that they wish to observe. It requires greater preparation time and special programming skills to establish the system, but these are available in a commercial package and once operational, the system is stable. Greater accuracy may be of benefit in playback also (Figure 3.11), though this awaits proof. Factors to be considered are summarised in Table 3.4.

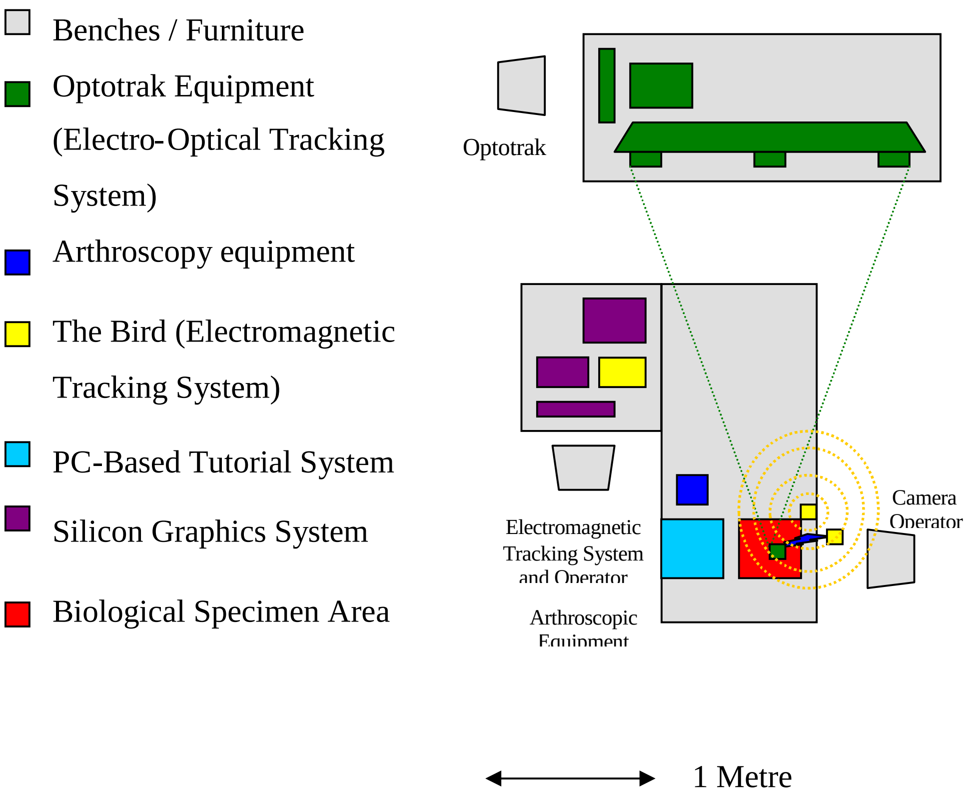


Figure 3.10 Laboratory set-up

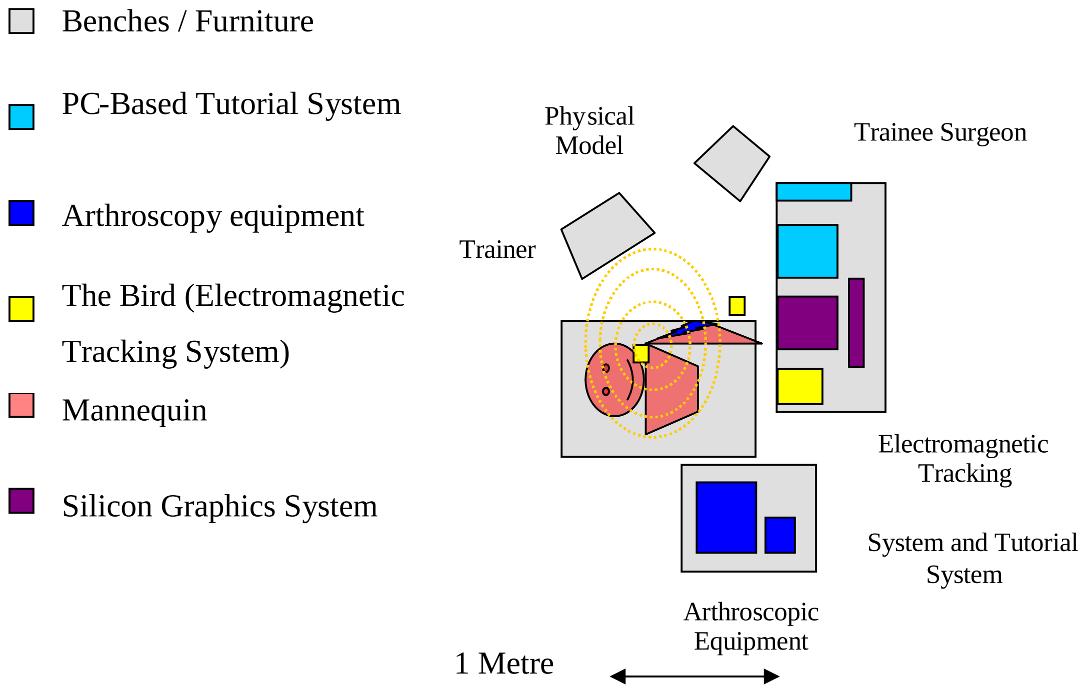


Figure 3.11 Proposed training (non-critical) environment using an electromagnetic tracking system

Features		**Optotrak™**	**Bird™**
Mechanism of performance		Electro-magnetic	Electro-optical
Scientific specifications	Effective range	2.2 – 6.0 metres	1 metre
	Accuracy	(x,y) 0.1 - 0.3mm (z)0.15 – 0.45mm	2.0 mm
Operating Location	Sterility	√	√
	Mobility	+	+++
	Fast set up		√
	Line of sight	√	
Cost (in year 2004)		~ £70,000	~ £4,000

Table 3.4 Comparison of electromagnetic and electro-optical systems for the preparation of biological databases.

3.14 Other computing issues

These are listed to provide tacit knowledge for those wishing to experience this domain.

Machine factors - bandwidth limitations - Processing capabilities reflect the potential conflict of focused resource versus distributed resource. Most hospitals now have T1 lines and with IPv6 a guarantee of service is achievable.

Sensors and actuators - The progressive development of appropriate input and output devices for VR simulation (73) requires advances in the allied technologies of sensors and actuators (74), as these progress from millimetric to micrometric and ultimately nanometric scales.

Speed of processing - There are certain applications for which the *Java* interface discussed in Chapter 4 is not the most appropriate tool. With the development of complex simulations, the use of *C* and *C++* for coding provides a fast and effective language. Designers need to keep this fact in mind when considering multi-user environments where critical action sequences must be maintained in the simulated environment.

3.15 Summary of surgical factors relating to distributed simulation in healthcare

Surgeons need to be able to adapt their environment to meet ever-changing needs. The user interface should be as simple as possible, acting as a bridge between experienced and non-experienced users. The basic technical knowledge of end-users needs to be addressed and this issue is taken up along with the educational issues related to 'distributed' simulation in healthcare in Chapter 5.

This chapter outlined issues of human performance and its limitations that need to be considered for this work. An analogy with aviation industry developments is used to emphasise the equivalent surgical attributes and sub-skills required. Problems of equipment design were highlighted by reference to equivalent issues experienced with cockpit displays. Poor human interfaces have been intentionally designed out of these systems. The intention is to do the same for surgical training. Attribute performance analysis is discussed as a basis for the AIPES surgical simulator validation system outlined in Chapter 6. By considering the issues surrounding the surgical navigation systems interface, the methodology (AIPES system) for assessing surgical simulators will need to take into consideration the whole scenario in which the simulator is employed. The surgical training environment and scenario factors where such knowledge (be it procedural or declarative) is required should be included to set the simulation in context for evaluation.

A specification for the training environment for shoulder arthroscopy pattern recognition skills was prepared based upon the user scenario profile identified by practising surgeons. Ultimately development of training systems employing simulation technologies requires a design process for the following three domains:

- Image collection in an actual operating environment
- Simulation in a training environment.
- Augmentation of the experience in the operating theatre environment

The technology of video was developed for the 1st EVW prototype, and a realistic target is set for distributed simulation, but the underlying problem is to ensure the specification is set accurately for the demand. Further development necessitated *‘down grading’* to 2 *dof* video worlds. This led to pervasive access by users and for technical factors of simulation to be incorporated into the evaluation process.

For this reason, the technical factors identified by Zeltzer (75) (AIP cube) are included in the AIPES scoring system detailed in the next chapter. The thesis aims are to improve HST training through application of virtual environments accessible over distributed architectures. The hard but realistic choice was therefore to compromise upon database complexity to ensure the distribution aspects could be managed successfully. As the opening paragraph stated, the thesis focuses more upon the human computer interface than the databases themselves.

4 *The evaluation of a shoulder simulator*

4.1 Overview

To assess the approaches adopted in surgical simulation, the simulation environment needs to be evaluated formally using a scoring system that considers more than just the virtual environment itself. The *Autonomy, Interactivity, Presence, Environment & Scenario* (AIPES) scoring system was created by the author and devised specifically for the purpose of surgical simulation evaluation. The results of an initial usability study, carried out to support the work and comparison with testing upon another simulator, provides a foundation for *Validation* and *Verification* in this field. It offers a framework for the evolution of formal evaluation methodologies that can be refined as feedback from users becomes available to close the audit loop, detailing the simulation's actual effectiveness, not just that implied by research evaluation studies.

4.2 Introduction

Two approaches to formal evaluation of the simulation environment have been adopted. Firstly users were asked to complete usability questionnaires for the evaluation of the hypermedia environments, and also the adaptive hypermedia techniques (76;77) that have been integrated into the core of a web-based learning environment. Results of this are reported later in Chapter 8. Secondly, to ensure that this did not adversely effect the simulation assessment, the simulation environment has been assessed separately using the AIPES scoring system. This is the purpose of this chapter.

The specifications of simulation systems thus depend upon how one is able to collect the data and make it readable to the user/viewer. This requires technological consideration of the user's ability, the resources available and the potential cost implication for the user or the institution. The ultimate aim is to set specifications including the provision of a baseline for development of new approaches to verify and validate the systems developed and evaluation of the progress specifically with development of the shoulder simulator.

AIPES provides a foundation for multifaceted evaluation. This considers the simulation, the educational environment and the surgical scenario involved in the training process. The evaluation of a simulator needs to be able to adequately justify its metrics. This reflects in part, the eventual 'usability' of such a system. Once the systems are introduced into the clinical environment, AIPES will stipulate specific qualities of a training environment, indicating its limitations.

AIPES attempts to encompass any simulation techniques or methods. It relates the fundamental features of a simulation (autonomy, interactivity and presence) with the context-dependent features relevant to distributed virtual university architectures. These include the learning environment and the educational scenario. The result of a usability study, which has been carried out in the third EVW prototype is described here to demonstrate the principle.

4.3 Setting standards for surgical performance

To break through the performance testing paradox[16] and establish internationally accepted standards, the author has adopted an approach based upon work arising from the 5th Cambridge Conference on Medical Education referred to by Newbel and Southgate (76;78). The sequential steps for assessing clinical competence are as firstly to determine the purpose – in this case, shoulder arthroscopy training, then define what is to be tested – e.g. orientation, navigation, pattern recognition. This leads to three steps;

1. Identify the desired clinical level of resolution. This provides the objectives.
2. For each problem, define the clinical tasks for which students are expected to be competent.
3. Prepare a blueprint to guide the selection of problems to be included in the assessment procedure.

[16] The paradox refers to the assertion that one cannot test the performance of a system unless the performance of the individuals testing it is already known. This cannot be clearly defined without knowing the performance of the tools used to test them.

4.3.1 Building a testing methodology

There are three steps to gain support of the end-user community:

1. Select test methods, which are most appropriate to the clinical task being assessed.
2. Let the clinical task dictate the method by which it is tested. This can be difficult to achieve.
3. Recognise the practical constraints on selecting optimal examination methods.

When addressing methods of testing administration and scoring one must decide on the level of efficiency needed in the particular testing environment. This includes how a student's performance is to be recorded or captured. It is then necessary to determine a method to assign scores for examined cases and all elements within the cases. Teams must take appropriate steps to ensure that the test provides unbiased measures of performance. They must evaluate the need for equating scores across different examinations. Finally they should review the procedure to ensure that it has not been trivial. When determining what is to be tested in these trials, it is necessary to elicit the boundary of competence. Relevant parameters are outlined below.

- **Identify the desired level of resolution:** This is designed for higher-level surgical trainees. During the first four years of their training, they are likely to complete a shorter module as part of their CCT preparation and this may also be used as part of the shoulder master class for HST years 5 – 6 in preparation for CCST. It is not currently considered as a potential tool for revalidation training. *i.e.* Consultant Grade., though the principles are the same.
- There are 3 levels of problems to resolve, starting with **orientation**, then **navigation** and then associated structure of **recognition** (leading to pathology recognition). It is necessary to identify the issues within each of these. This has been developed as part of ergonomics analysis of the system, in conjunction with the Helmholz Institute in Aachen. The key issues are the frequency of problems and the frequency of occurrence, importance, and severity. The shoulder simulation has been developed to model a procedure that is difficult for trainees to view, as it occurs infrequently, but is becoming more common. Like any arthroscopic surgery, it is important as it offers a potential cure for painful conditions and amelioration of others, and also carries the risk of severe complications, such as vascular or neurological injury.

- **Finding the clinical tasks within each problem** Specific terms are found, and the level of resolution appropriate to the expected performance of the trainee is described. The definitions, boundaries of accepted thresholds for normal and abnormal and also specified clinical management paths (algorithms[17]) are identified. To save time, this needs to be pragmatic.
- **Preparing a blueprint to guide task selection** Comprehensive content and competencies blueprints (checklists) are developed for the assessment procedure. This generates a multi-dimensional "grid" which considers the various factors to be evaluated relating problems to categories of competence. The blueprints indicate specific and critical tasks embedded within the problems which need to be tested. This forms the basis upon which the sample for testing is selected which, although ideally random, is, of course, going to be limited by the number of those able to test the system. For the assessment procedure, a high level of content validity should be assured. Once implemented this will rely upon 'expert committees' who review the material.

4.3.2 Selecting test methods

Although a wide range of methods for assessing clinical competence exist (79), only some of these methods can be appropriated for the discrete tasks that can be performed using the simulator. The three steps are outlined below:

1. **Selection of the most appropriate methods for the clinical task to be assessed.** Where possible, the questions posed to candidates are structured to allow didactic diagnostic yes/no answers posed. With the assessment of clinical skills, validity can only be achieved using multiple observations and performance.
2. **The clinical task dictates the method by which it is tested.** This returns to the principle that what is easy to test is not necessarily useful and what needs to be tested is not necessarily easy to test. No single method is therefore capable of measuring all components as more than one test method is used within the simulator assessment procedure.
3. **Practical constraints upon selecting optimum test methods.** Clearly, this level of generation of the simulator can only test certain aspects of the simulation, hence the concept of building in scenarios. As a consequence, only the validation and

[17] Examples of Medical Algoroithms can be found at www.medal.org

verification (V&V) procedure for teaching and testing these discrete skills can be used to evaluate it.

4.3.3 Addressing the issues of test administration scoring

There are 6 steps to follow, with respect to the administration of scoring:

1. **Decide the required level of efficiency** As mentioned above, more than one test is required and these ideally should be arranged hierarchically so that the most efficient test is the first administered. It is, however, important that the tests do not just reliably discriminate between candidates but have their purpose based upon discriminating clinical competence. This returns to the issue of formative versus summative approaches. No operation is always a success, since every procedure carries with it the risk of complication and failure. The approach is one of providing an average percentage risk as part of informed consent and as Poloniechi (76;80;81) points out, half the surgeons are, by definition, below the average results because that is characteristic of a normal distribution. One surgeon, of course, will have the worst result. The aim of the assessment is to demonstrate that the skills have been acquired in all the necessary competencies and such skills are likely to be explicit as Klein (76;82) suggests. Ideally, this will ultimately progress by adaptive testing, where each component of the assessment is selected on the basis of performance of previous components, although the complexity of this strategy needs to be built within the overall model. An alternative strategy would be that of sequential evaluation, which is more akin to the gaming model of achieving certain levels of performance and then being tested upon the next level, if they are successful in the previous one.

2. **Decide how performance is measured and captured:** Issues are raised such as security, training, and consistency between sites and thus the centralised model of client-server architecture was developed. Pilot studies used a pen and paper approach. This method, however, is not viable for the progressive studies, for which an automated process was developed. It is necessary to introduce computer-based testing multi-site, multi-station systems, for examination and this led to the drive to develop the *Virtual Orthopaedic European University* (VOEU) project detailed in Chapter 6.

3. **Determining the methods of scoring the cases and the tasks within the cases** The case refers to a conceptual unit for assessment. It may not necessarily be an individual clinical case. With the development of the system to incorporate different pathologies, this is likely to become a suitable level for structuring the database, since it

also allows for factors such as authorisation of use of clinical material, *i.e.,* consent. Once the case scores are obtained, then the process of combining these for decision-making becomes all-important. The underlying principle is that, by combining the procedural experience of using the simulator with the form assessment, there is added educational value in doing the test. Unsuccessful sections can be repeated until the score demonstrates the pass / fail level of competence.

4. **Steps taken to avoid bias** This is likely to be introduced when a minimum threshold is being set. In the earlier models, 'time' bias (generating pressure to complete tasks within a certain limited time) can be introduced into the simulation. The only reliable method to improve the statistics and overcome this is by observing a wide sample of performance, recording as many users as possible with an independent marking system. This ongoing recording system allowed the numbers of users to be increased on an '*ad hoc'* basis, so as to allow for asynchronous collection of results data, minimising the time pressures upon the user 'test' population and to allow statistical analysis upon larger populations (76;83;84).

5. **Evaluating the need to retain scores between examinations** Within the scope of this thesis, the aforementioned design for examining considers the simulation in isolation. It does, however, relate it to the AIPES system referred to in detail below. It is through this framework that it may be possible to equate different simulation systems for their validity in performing the roles required. There is, however, the risk here of comparing apples with oranges, since simulators will be designed to test different systems for different purposes. The important correlation is between performing the same test upon different systems at different sites to ensure an accurate correlation. Alternatively the candidate should know to fix upon certain points or pathologies, which are adequately represented with different systems. This 'anchor test' could be introduced as part of the registration process for starting a new test. This effectively represents a generic *'educational scenario'*.

6. **Ensuring that the test is not trivialised** This depends upon scoring the relative importance of various components of the test with appropriate weighting. This is only significant when the test becomes part of a formal evaluation system such as the CCST.

4.4 Setting standards for surgical performance using EVW prototype one

The previous section leads to the question of setting appropriate standards and the process to achieve this. The score to pass the examination needs to be widely accepted. Such procedures are, to some extent, arbitrary but should not be capricious. Referring to the 2 steps involved in setting the standards:

1. Determine the type of standard desired and an appropriate standard setting method. For the shoulder arthroscopy pattern recognition sub skill standard users should be able to navigate around a joint, and identify sites of common pathology – tested upon a population of users.
2. Develop procedure for effective communication of results of the test. This is managed by the dynamic review journal in the Virtual Orthopaedic European University (Chapter 7).

It is from this approach that a core methodology can be constructed for the preparation of simulation evaluation. A simulator can only provide an accurate model if one has already decided what "accurate" actually is. Some would say 0.1 mm actual resolution is necessary (69) equivalent to about the width of 10 average cells, whereas others would accept wider tolerance. Darzi *et al* consider that "finding objective criteria for judging good surgical technique is difficult" (85) and Stone goes on to suggest that the word "good" should be replaced with "any" (71), but the systems used to obtain quantitative measurements of performance. As Stone points out, reliable testing methods have been developed for assessing mental workload, cognitive performance, perceptual skills and situational awareness as integrated with appropriate experimental designs and regimes (36;62;86;87).

4.5 Determining the type of standard required

Most examination systems adopt relative standards, which are based upon a statistic where a bottom percentage has failed. This, in effect, means that one reverts to a normative system of valuation. This is not really appropriate for a test of clinical competence. Absolute standards are based upon the analysis of the content of the test and this is based upon the judgement of experts as to the expected scoring for a borderline candidate. Since this can be related to the legal standpoint upon medical competency, it is perhaps a more reliable system than at present. One alternative to this is to produce a reference group. This is the principle underlying the simulation system, since, as a broader database is collected of user performance results, this can then be used to provide an accurate correlation between a set of normative values and the expert judgement. Ultimately, a 'closed loop' could see novices themselves become experts and a minimum standard developed. The method is likely to entail a large sample of candidates, which can only be achieved with multi-centre trials. Their performance data can then be used to develop a minimum standard. This does impose data collection and analysis requirements to support the process.

4.5.1 Developing the procedures for communicating the results.

This process requires both peer review for validation and the appropriate method for dissemination. To this end, the VOEU project has been developed to meet this demand, with inclusion of a Dynamic Review Journal (DRJ) component. This is explored in more detail in Chapter 7. There are other independent international standards organisations which will be involved in future work. The principle of the DRJ is that the procedure becomes transparent, both with respect to the assessment procedures and the decision-making process. The evaluation can be made available on-line form (document) for candidates to review.

4.6 Determining the purpose of integrated simulation

There needs to be a clear purpose for the assessment of the simulator itself. In the case of the procedural training simulator, this should be to demonstrate competence of an individual in their ability to orientate, navigate, and recognise certain structures within the shoulder joint. It is for this reason that a *virtual reality modelling system* was chosen rather than, for example, solely objective tests such as an MCQ. There is however the need to use an MCQ type of system to collect the results from users in the trials.

This must comply with the overarching pedagogy accommodating different learning resources, such as the simulation systems that must clearly state the aims, objectives, and process of evaluation as part of its overall specification. For the demonstration system proposed, the author would suggest that the system is integrated into a problem-based learning scenario to add realism to the model's training environment and this is explored in more detail in Chapter 5. This should add a sense of realism as surgeons can more readily relate the questions to their experiences.

4.7 Validation phase 1 methodology for simulator assessment

Assessment of surgical performance is a moot point. It is not only a final and evaluative part of the instructional unit. The assessment process occurs throughout the activity. The situation is complicated by the user's need to perform clinical duties. Clinicians were only available for one-to-one demonstration and review of the system, with either on-line or paper-based questionnaires to complete. The solution to this was to recruit a trial manager, Dr Craig Reston MRCS, who was able to dedicate the time to the individual user within their work schedule.

Typical virtual environment topics for evaluation: This method may be used to assess any system, from a stand-alone desktop system that is designed to train a predominantly visual skill, to one based upon a distributed architecture as demonstrated here. This can be evaluated when performing the necessary skills of surgery. The key issue is not whether to creep forward through evolution of digital substitutes for other education, training and archive technologies, but whether to promote the revolution of clinical practice through the integration of pervasive computing technologies. By using both 'Just-in-time' as well as 'Just In Case' educational strategies, providing tools that deliver anonymised patient-specific data to the user in a virtual environment, it is potentially possible to seamlessly integrate this by virtue of an educational pedagogy. People need to strive for the vision (61).

Such early prototypes are largely being superseded; however, virtual environments can be updated easily and automatically. The move is toward *autonomous computing*. The use of a gaming on-line environment can be employed to allow users to run rendered wire frame models on their local machines but the other components can be delivered by store-and-forward services prior to the start of the training session. The opportunity is also there for feedback upon different versions of software, using such systems.

4.7.1 Educational environment generation for evaluation

This needs to meet the basic educational demands of instructional design with clearly stated goals and objectives. The methods of feedback and remediation are of interest as they pose more technical demands than the other aspects that are predominantly text-based. The following issues were addressed:

1. **Conduct assessment of the trainee performance:** in terms of a quantitative notation. Ideally this will model the actual track of the user for replay and discussion, but this is beyond the capabilities of the third generation EVW prototype.

2. **Provision of feedback:** right/wrong feedback with justification and relation to the subjacent theory and/or recall of the related theory and/or the visualisation of the result of the user's process. At present the aforementioned supervisor provides this feedback. This will ultimately be based upon the user tracks referred to above.

3. **Provision for remediation:** redirection to other steps of the process – This is managed by bidirectionality of the streamed video.

4. **Redirection to a prerequisite section** or recall of the related theory – This is managed by links using hypertext links embedded within the video sequence, directly linking to the library components.

4.7.2 Internal structure and content of the educational scenario

This is the framework for constructing Multimedia Educational Orthopaedic Module checklists. These may be added by a tutor, as in the test scenario, when issuing the directive reading, with explanatory text for that specific course or outlook upon the generic text of the handbook. The handbook provided to the trainees does not closely follow all the guidelines of the surgical course model. The guidelines are in accordance with the structure, though the 'trainee' requirements section is not. Evaluation of the instructional event for the trial relating to the arthroscopic shoulder operation is found at: http://voeu.ecs.soton.ac.uk/ : VOEU : Library : Shoulder : Instruction : in three documents entitled: introduction, body, conclusion. The aims of these sections are:

Introduction

The purpose is to inform trainees of the purpose of the instructional unit, and of the professional interest of this. It should provide overview and description of how to use the instructional unit. It also provides explicit required background for user profile (professional issues) and prerequisites (knowledge issues).

Body

Information is presented in accordance with the chronological structure of the presented surgical intervention. At each step of the chronological structure information is presented according to the Mayer's Multimedia Principle (88).

- Present information in different settings using text and diagrams or illustrations to detail the equipment, anatomy & recall related theory.
- Provide interactivity, *i.e.* opportunities for the trainee to test out and try understandings in some way – to act using gestural (gesticulation or dexterity)[18] and intellectual skills
- Provide feedback to the trainee's actions: This includes explicit expert's criteria for validation ("how can the surgeon decide that each step was achieved correctly"), by presenting information in different settings. Explicit specific terminology (definitions) should be described and right/wrong feedback given to responses, training users to recall related theory. This will provide a different setting for viewing the result of the user's action.

Conclusion

This provides a summary of the instructional unit in terms of learning objectives "now you should be able to..." and provides further reading, whilst restating motivation aims. To adapt this generic approach for the shoulder module scenario, it is possible to build a specific checklist for the purpose. The principles should however be adhered to for comparison between simulation types.

4.7.3 Technical requirements

The following requirements were met for local alpha-testing by providing a catalogue of key components and requirements. This included a browser add-on requirement list, and a list of files and functions needed to support the system component. The *Reviewers Guide* has evolved during the beta-testing into the *'Help Guide'* linking into the *ICT* handbook. An underpinning RCS Educational Strategy has been written to provide a foundation for this.

Factors used to evaluate the technical system act as instruments for developing future surgical educational systems. They can be grouped as:

- Assigning points to effort – Extension of the eSTEP scheme
- Time vs. effort need to be defined for certain subtasks

[18] The first two steps (equipment and anatomy) are in place, but for the reasons given in section 1, dexterity and intellectual skills are not included. Again these can be added by the clinical tutor when setting the directive reading.

- Ratio (balance) of effort upon different subtasks is set up by the individual – each person can contribute material to the educational resources, i.e. 'own a module'
- Length of time users spend upon the creation of the material
- Active as well as passive learning approaches
- Weighting system for time and value of effort expended on tasks and subtasks by Stakeholders
- Feedback upon values of contribution from various authors

4.7.4 Scenario generation for evaluation - Relating the educational scenario to the pedagogy

This relates to the user *questionnaire*. The following pedagogical principles need to be addressed and are scored on a 5 point 'visual analogue scale' (VAS):

- Feedback during the activity (a series of steps) that any actions or thoughts were either right or wrong
- Opportunity for the problem solver to start a single step in an activity again
- Feedback during a single step, will not prevent the problem-solver from accessing the following steps of the activity
- Thresholding of the results so that the problem-solver cannot proceed to the higher-level activities until he has successfully completed the lower level activities.
- The information that the problem-solver must be given, obtained from the investigation over time, allowing reasoning to occur at every step of the activity (89).

4.8 Shoulder simulator evaluation methodology

To approach the resolution of this performance testing paradox (page 44), two groups of surgeons and tools need to be separated from the main results of data collection by considering firstly the ideal purpose of the simulation system and then looking at how it can achieve these goals. To do this, the approach involves the surgeon performing a task and subtasks that are allocated unique identities. It is necessary to decide which subtasks are identified as unique events. A reliable way of tracking both the actions and which variables are measured (e.g. visual and haptic) to develop a 'ground truth' is necessary to record the performance of the individuals.

This represents the formation of a sound evidence base, though it is resolving this and developing methods of simulator testing which may then be used to test other systems that is of paramount importance to the field of Medical Education and so a generic methodology for evaluation of such systems is proposed as part future methodology development work.

The calibration / evaluation of such a system necessitates the simulator being operated (tested) by a significant population of experienced surgeons who have completed the learning curve and are able to trial the system. This again fits well with the aim of such systems being first assessed by a test surgeon who has expert experience of how the real environment of the surgical procedure behaves and thus is able to compare and assess the simulator with regards to the degree of realism. This allows for iterative design of equipment without relying upon clinical beta-testing, avoiding the risk of direct consequences, and the potential ethical dilemma this would create.

4.8.1 Approach to evaluation specification development

Integrated Simulation incorporates both declarative and procedural knowledge, providing trainees with a chance to acquire knowledge upon orthopaedic disorders as well as develop focused clinical reasoning skills through a series of graded case studies. In order to enhance learning and reduce cognitive overload by developing a new user interface for surgeons that adopts a form of surgical simulation of an operative environment, simulation has been employed for the modelling of shoulder arthroscopy. Two methods were adopted to formally evaluate this novel technique.

- Users were asked to complete usability questionnaires for the evaluation of the hypermedia environments, and adaptive hypermedia techniques (77) that have been integrated into the core of a web-based learning environment.
- The simulation environment has been assessed separately using the AIPES scoring system.

4.9 Verification and Validation (V&V) process

This is defined as testing a system to check that it is suitable for the tasks intended. To address this question, the main study required the population to validate as well as verify during the evaluation. This demonstrated feasibility by collecting the user results using a paper-based form (questionnaire) and integrating them into data collected using the purpose-built dynamic review process of the VOEU site. This supports trials management and statistical analysis. It is yet to support direct generation of the formulae specified above, but the author concludes that this can be incorporated if the validation process recommends it as part of ongoing on-line evaluation of simulation systems (Figure 4.1).

4.10 Autonomy Interactivity Presence, Environment and Scenario (AIPES) methodology for surgical simulators assessment

4.10.1 Application of AIPES - clinical sub-skill modelling for validation

There are two basic ways to approach a problem: often referred to as ‘top down’ (a reductionist approach) and ‘bottom up’ (a constructivist approach). These methods have been used in various ways to address the issue of how to assess simulation for surgery. Table 4.1 segregates the commonly used methods according to these groups.

These will be detailed in later sections but the major transition is from established reductionist approaches adopted to analyse the initial evaluation problem to the specific constructionist analysis tools developed in conjunction with the Helmholtz Institute Aachen (HIA Germany) to build the evaluation protocol adopted for the main study.

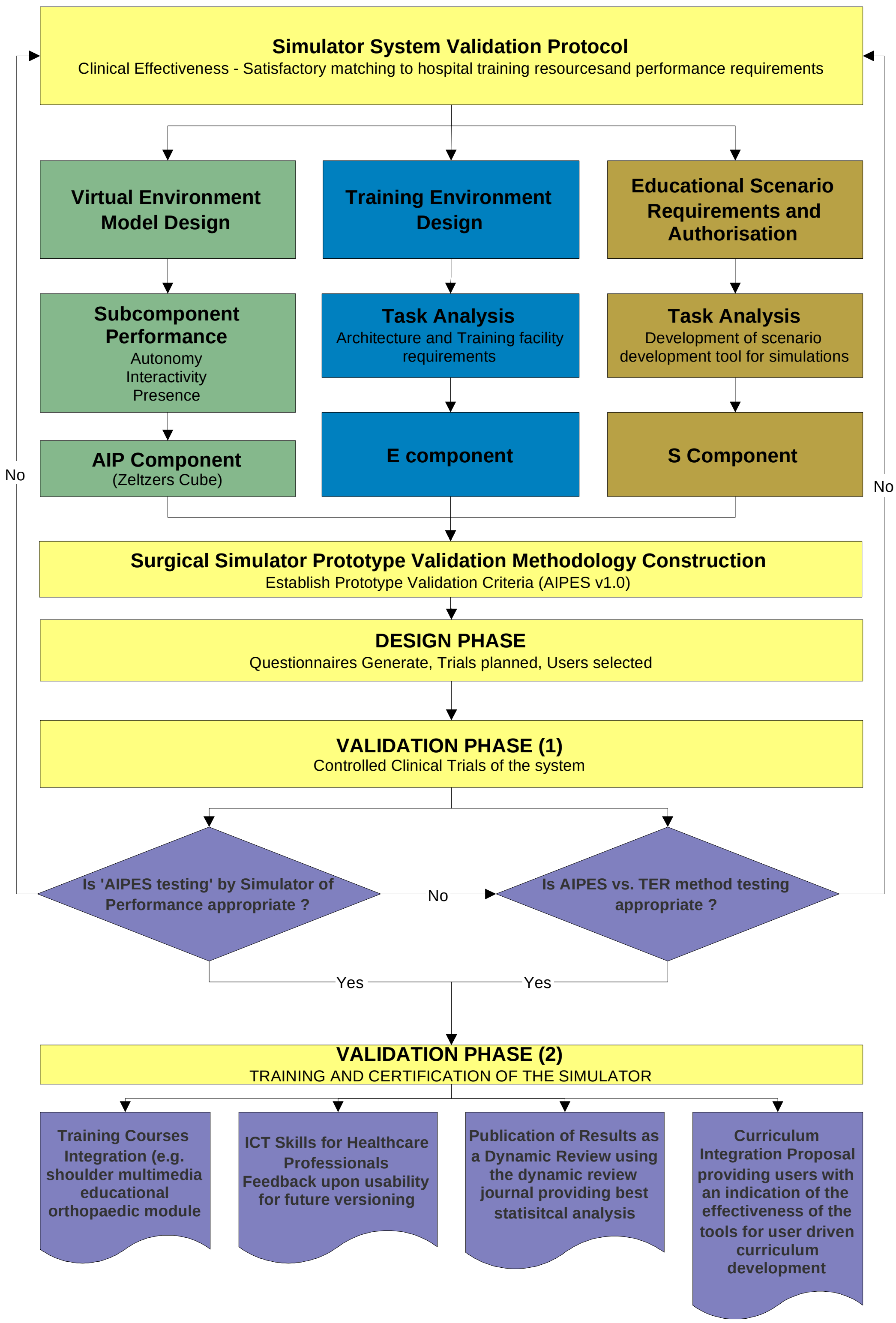


Figure 4.1 Modified Validation and Verification process [V&V] based upon Higgins et al, modified to show integration with curriculum development, so that planned embedded components can be integrated into the surgical curriculum. This leads to the process of embedding the design of simulations into the curriculum also through the iterative process.

Reductionist	Constructivist
AIPES	Task Analysis
NLP	Sub Skill component clusters

Table 4.1 Constructivist vs. reductionist approaches to simulation evaluation

4.10.2 Reductionist tools developed for EVW analysis – AIPES

The AIPES measure of performance of the simulation system is derived from the following evaluation measures detailed elsewhere (90):

1. **Autonomy:** The ability of the model to provide realistic actions with or without response to stimuli without input from the user measured on an analog scale by users (Appendix 4.1).
2. **Interactivity:** The ability of the model to provide realistic actions in response to stimuli input by the user.
3. **Presence:** The ability of the model to convince the user that the environment is realistic.
4. **Environment:** The ability of the model to provide realistic environment as compared to real life surroundings as perceived by the user.
5. **Educational scenario:** The ability of the model to provide realistic clinical educational scenarios as compared to real life events and learning opportunities as perceived by the user.

Although the reductionist approaches were adopted for the management of these cases, the initial results suggest that this may be of value for independent assessment of the system, whilst the demonstration of adequacy of function will require a constructivist approach, and this has been adopted in later versions of the iterative design process. It would be of interest to the author to see at which stage the reductionist approach indicates that the expectations of the end-users have been met, and whether this relates to the design specification drawn up using the constructivist approach. Future correlation between the user design brief and constructivist design specification could be evaluated objectively by the end-users. Using this HCI perspective, future adjustment consequent upon the ‘reality index’ values outlined above will be made. These are set to parity for the purposes of the pilot study as in Table 4.4.

The AIPES system is a simulation evaluation tool, which refers to the autonomy, interactivity, presence, environment and scenario aspects of the simulation and its educational potential. This adopts a top-down approach to solving the problem of what needs to be evaluated rather than what is straightforward to evaluate. It is therefore accommodated into the scoring system but will not be fully exploited until implementation. This raises two questions that should be answered:

1. Should the sum, product or the mean of the factors be used?

Taking the sum would have better reflected the views received as part of the feedback from the people using the system. The sum comes out at around 50% effective and makes the simulator appear to be better than the product of the variables for the score, which accentuates the lowest score. The aim is to produce a discrete comparable value consolidating the five separate variables. The intention is that AIPES could be adopted as a mean of the individual scores (factored and then converted into a mean). It would also be possible to express AIPES as a vector by using the square root of sum of the different factors, which would aid its visualization. This approach was not used during the first three rounds of evaluation as the comparison with the knee simulation from SSSA (page 91) used the product, though the author will employ the mean in future trials.

2. Is there the need for an overall weighting - 'c' coefficient?

This is intended to weight all the factors rather than just the AIP factors only (as per the name). This is demonstrated in the example. This takes into consideration the time actually available to train individuals for the tasks. C_A represents the coefficient for the ratio of time spent in simulation exercises whilst training (e.g. attending a course), compared to $C_{B,}$ the time spent in operative exercises during a training post. For the initial trials these are set to parity.

4.10.3 Principles demonstrated by the AIPES scoring system

- It considers the aspects of VR, namely *Autonomy, Interactivity and Presence*, but also the relative influence of these collectively, and the environment in which the simulation is set. This environment is both conceptual and physical. One needs to consider the training laboratory factors, *influencing* the work performed and also the generated scenarios that can be used for training.
- The underlying concept is that the simulation should be scored comparative to a real surgical procedure with respect to its accuracy and value as a training resource for

discrete surgical skills. This leads to the principle of an **A-B comparison**, where A is the simulation and B the real procedural exercise. As such B may be part of an operation or the whole procedure. This should be part of the specification for the procedure and the training resource.

- The A-B comparison implies that the simulation would ideally be as good as the 'real thing' i.e. learning the same procedural technique in an actual operation. This is not so. In fact it could potentially improve upon certain surgical training exposure, and of course it eliminates risk to the patient directly. Since a more effective training tool may be generated in the future it is sensible to assume such progress, and accommodate it in the scoring system. The correlation factors for actual surgical practice may be less than parity as the surgeon cannot necessarily 'choose' the ideal case upon which to learn and the actual time when operating may only be between 10-30% of that perceived as most of the time spent during an operating list may be of little educational value.
- The system is discrete, independent of other variables that need evaluation.
- It compares the simulator to the skills / functions for which it was designed, rather than a generic testing system.
- A single value is produced to facilitate comparison of generations / versions of the simulator. This is intended to provide a relative measurement (ordinal scale) rather than an absolute one.
- The assessment of the system is multimodal, reflecting the different sensory modes of interaction mimicked in the simulation environment.

4.10.4 Adjusting coefficients

To compare the relative merits of various components of the system, coefficients are included to allow for likely modifications. This is the first of such systems used to assess / evaluate surgical simulators and is thus likely to undergo much modification. It is applied in Chapter 6 on the pilot study of the user perception of the relative significance of the modalities.

4.10.5 Visual feedback – interactivity

The 'quantum leap' in surgical simulation was the development of 'Real Worlds' representing the body cavity, which was being navigated. Such worlds are necessarily very complex: the anatomy must be represented in three dimensions and must be accurate. The view through the virtual endoscope (usually offset by 30° and magnified) must change instantaneously as the endoscope moves. The rate of change must be more than 20 frames per second in order to generate the impression of moving images of a real world rather than a series of still pictures. Finally the colour and textures of this real world must be sufficiently realistic to convince the operator that he is operating in a real body cavity and not a computer simulation. There is a spectrum of methods by which such complex *'worlds'* can be simulated. This ranges from a rendered *‘graphics world’* constructed upon a wire frame model, to a 'video world' such as EVW. The ultimate aim of convergent design is to achieve the features of both whilst minimising the drawbacks. This is discussed in more detail in Chapter 9.

In the EVW Prototype 1 a three-dimensional wire framework of the body cavity is generated in the computer, and then this framework is rendered to look like the body cavity. The Prosolvia group, now relaunched by Prodiscus (91) has developed a shoulder arthroscopy surgical simulator, which allows triangulation and tissue manipulation to be performed within a graphics rendered world. However the problem with this form of simulation lies in the accuracy of the rendering (92;93), for currently such systems still visually resemble a plastic model.

4.10.6 Haptic (force) feedback

Computer systems are now being developed for virtual reality, which provide tactile feedback from the world to the operator's fingers. The *TELEOS virtual knee7*™ apparently provides realistic force feedback depending upon which virtual anatomical structure is in contact with the virtual instrument, simulating the give of softer tissues and the resistance of harder tissues.

4.11 Evaluation of subtasks using the AIPES system

One can break down individual subtasks into the list of necessary surgical skills that can be taught by surgical simulation using VR technology. At least 100,000 concepts are known and documented in medicine, with no clear ontology linking them. There is a list of necessary tasks, which for shoulder arthroscopy includes orientation, navigation, and pattern recognition. As with *‘aviate, navigate, communicate’*, these are prioritised. Each is important. It is unlikely that the system will display a linear relationship between different characteristics. There is no direct relationship between sense of *Presence* and the degrees of *interactivity*, nor that will either of these correlate directly to the *autonomy* of the model. There must therefore be two further levels of complexity. Primarily there will be some relationship between *Interactivity*, *Presence* and *Autonomy*, indeed all three together.

This was first described as 'Zeltzer's Cube (Figure 4.2) in 1992 by Zeltzer (75) and used as a way of describing the 3-D vector for realism of simulation. Secondary factors involved include the "realism" of the surrounding training environment e.g., the physical equipment used. These factors are harder to quantify, but will have an effect. The relationship between the individual sub-components and the scenario that is generated for training is important here. These may not be related directly to the VR components of the system. They are restricted by its performance. There is also likely to be the secondary effect, having not just primary relationships between the three axes of VE performance, but also secondary effects upon each other since the autonomy and interactivity will affect presence.

To stretch the aviation metaphor further, an analogy with this is the controls of an aeroplane. The rudder, ailerons and elevator will all have different characteristics of force feedback, so far as the operator is concerned. These may not only provide characteristics discernable so far as the operator responds to each control, but also the general feel of the aeroplane as a whole. Thus there is a "comfortable ratio" for using these different controls. These aspects of aviation have to be managed along with the navigation and communication by the pilot. Such a balance may exist for surgical simulation. The analogy fits with the additional tasks generated in scenarios.

Fidelity is not just the virtual environment (VE) and the virtual reality features that this generates. There is also the way that this resource has been set up in the context of an integrated training system. One needs to consider the training ***Environment*** in these desktop systems, and also the exercises that are trained – *i.e.* the ***Scenarios*** generated. This supports the educational concepts of the Problem-Based Learning (PBL) process. It is outcome driven, based upon the curriculum seen in Figure 5.1 below.

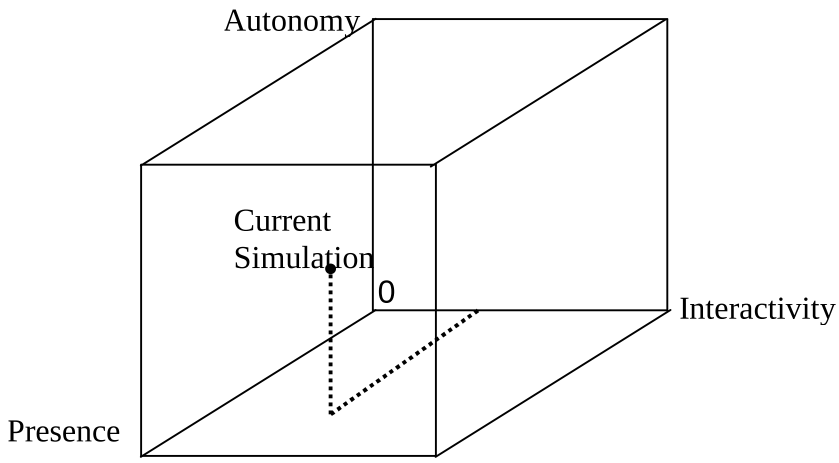


Figure 4.2 The relationship of autonomy, interactivity and presence, after Zeltzer's Cube (1992)

4.12 Suitable methods for VE evaluation

This formula describes the A/B comparison where A is the virtual model and B is the real life learning opportunity.

$$\frac{C_A\,((a.A_A) + (i.I_A) + (p.P_A) + (e.E_A) + (s.S_A)) \times 100}{C_B\,((a.A_B) + (i.I_B) + (p.P_B) + (e.E_B) + (s.S_B))} = \text{'AIPES' \%}$$

Equation 4.1 AIPES Formula

This may best be represented mathematically, as in the formula above with the following key;

A	=	Autonomy,	**a**= Autonomy Coefficient
I	=	Interactivity,	**i**= Interactivity Coefficient
P	=	Presence,	**p**= Presence Coefficient
E	=	Environment,	**e**= Environment Coefficient
S	=	Scenario,	**s**= Scenario Coefficient
			C= Collective AIP Coefficient

In the evaluation for alpha testing all coefficients were equalled to 1. The issue is how much of the discrete skill base is actually being modelled and if these skills are important should they be prioritised. This would benefit from comparison to other such models of skills training such as the *Objective Structured Assessment of Technical Skills (OSATS)*[19].

Improving the realism of any of these areas should be of benefit overall, but each may not be matched to the expectations of the users. The relative merits of each aspect of reality modelling are given different weightings, currently set to parity (1:1).[20] These may be influenced by the population perception of the relative importance of these factors. These may then be adjusted to provide an absolute value, by which advances can be judged. *The aim is to ensure a mechanism for validation and verification as outlined in* Figure 4.1. This should be based upon actual usability and user subjective evaluation. The alternative is a set of technical specifications, which may be of value in demonstrating the verification of performance parameters outlined in the design brief. Such an approach does not necessarily reflect the true 'value' of the system in the user's perspective.

Whilst the AIPES scoring system is crude, if accepted for the oversimplification that it is, it will allow comparison as various systems for their individual merits, providing equivalents. It provides the foundation for a '*protostandard*'. The different factors are added, to provide a single value, which can be expressed as a percentage, indicating a relative simulator performance and compared against clearly defined educational objectives. The critical issue is to compare the intended educational objective with the perceived ability of the simulator to deliver this outcome and to assess the effectiveness of the simulator for this task. This has been described in other terms in the vehicle simulation field. The term *Transfer Effectiveness Ratio* (TER - see below) has been used to reflect the educational value of the simulation systems (94). The TER value for each system should be calculated, in order to attempt to quantify its cost-effectiveness, but the relative potential of such a system is more easily measured than the true value.

Since it is necessary to start somewhere with the disentanglement of the multiple threads that make up an effective simulation and evaluation thereof, the questions posed in the pilot study do not necessarily portray the values precisely.

---

[19] More information regarding the Canadian OSATS programme can be found at; http://www.aaos.org/wordhtml/bulletin/jun99/cme1.htm

[20] All coefficients are initially set to 1.0 so as not to affect the scoring. They can then be modified as data becomes available via the virtual observatory (VO) and the dynamic review journal (DRJ).

4.12.1 The role of clinical controlled trials

The aim of the AIPES scoring system in gaining meaningful results is to be able to say whether a simulator is of value in training and to quantify its potential. The VOEU clinical user evaluation trials underpin the Western World's logical approach to building a framework of *Evidence-Based Medicine*. It is not however going to be practical to formally test every simulator for every sub skill on an adequate population of suitable trainees, for every version and update. A compromise has to be achieved. It might be better to establish a reliable training assessment protocol that may then be applied to new systems to ensure endorsement. At least one controlled trial will however be necessary to establish the precedent. This should in effect calibrate the protocol. This involves the VOEU project and is to be found in detail at the Southampton VOEU site in the DRJ (95;96).

4.13 AIPES results of the shoulder arthroscopy simulation study

To complete the upgrade of the AIPES v1.0 scoring system questionnaire for the preparation of the main study it was necessary to agree standard questions that are representative of the issues of the environment in which the simulation is presented and of the scenarios that it generates. The questions proposed for this are outlined in *Appendix 4.1*. These are designed to accommodate the different potential uses of the simulators in accordance with the *protostandard*.

In order to get agreement upon the questions, they were presented to the simulation committee of the VOEU organisation and take into consideration the effort of the VOEU education committee with respect to the assurance that simulations are indeed meeting the requirements of their design by meeting the educational objectives for which they have been designed. In order to facilitate the changing of these roles and the ability to reflect different educational scenarios in future, using the same generic simulation modules, the trial has been constructed with the questionnaire for the scenario evaluation as a separate form *(VOEU Questionnaire D)* that can then be mapped to the actual simulation scenario as new ones are prepared. This will allow for reliable feedback upon versioning.

The next stage of evaluation is of course the environment. This is mapped to the centre format for the trial. This may require specialist equipment and so in the first version of this, the format is the same as that for any usability analysis of a new system. In order to make this useful for the future development of VOEU and the simulation subsystems, it has been reclassified as AIPES v2.0 reflecting the major changes with the introduction to other developers and for different simulation systems. This represents the paper forms and so the digital on-line forms (kindly prepared by Dr Chris Bailey) are referred to as AIPES v2.1 in the metadata. The result is a relationship between different metrics that is shown in Figure 4.3. This shows how AIPES, through the development of the A/B comparison could provide a foundation for a *'reality index'* measurement (multiplying by 20 to convert the score into a percentage) for the simulator that can then be related to the educational potential as expressed by the transfer effectiveness ratio detailed below.

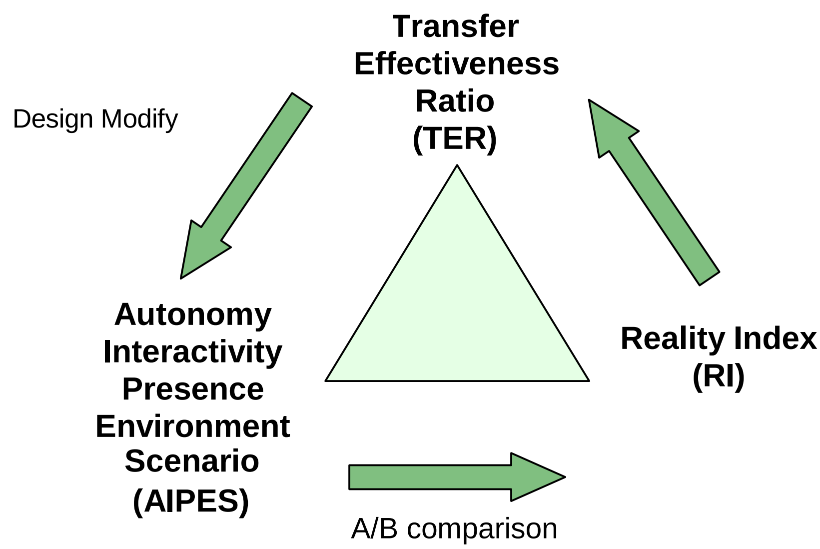


Figure 4.3 Relationships between TER, AIPES and the Reality Index

From an EVW prototype 3 experiment where A=0.74, I=0.65, P=0.62, E=0.64 and S=0.67 so Σ =3.3, the following pilot study results are obtained;

$$\frac{20\ AIPES_A}{20\ AIPES_B} = 66\%$$

Equation 4.5 Results of the AIPES score for the VOU3 experiment

This study of 17 trainees gives us an AIPES score of **66%** (66.4%) for the Shoulder Simulator suggests that the simulation is about 2/3 as effective as real life training opportunities. This is perhaps a realistic figure for the performance of EVW prototype 3. This is not proportionate to the range of answers one would expect when one reads the comments made by those using the system. This does fit well with the results of a similar trial performed lately by our Italian collaborators with their knee simulation system producing an AIPES score of **67%** (66.9%). Incidentally the EVW prototype 1 produced results between 2 and 17% effectiveness for the simulation alone. The improvement in interface could to a great extent be argued as a reason for the significant improvement.

This serves as a humbling reminder of how much development is required to achieve a realistic Virtual Environment (VE). It is the first step to elucidating the system of assessment for surgical simulators, providing a process for the advancement of this field.

4.14 The role of the Transfer Effectiveness Ratio (TER)

Returning to the question of the actual value of such systems as a way of comparing the ‘proxy’ value of the AIPES score with the TER, this adequately reflects the user perception of the value of the system when compared to actual operating time. This means that the valuation methodology has completed a loop of feedback and so can become part of an iterative design process for the development of the next generation of systems.

Traditionally vehicle simulators such as flight simulators have been assessed using a system known as the transfer effectiveness ratio (37) where the time to transfer skills to pilots using simulation is compared to that of real flying. This is expressed as a ratio that represents the equivalent real flying time to gain the same experience as one hour of simulation time. This has allowed for the review of a system indicating its ability to train an individual in a certain time frame to attain certain skill levels. An example is that for commercial flight simulators, a TER of 0.48 (57) which is widely accepted, suggests that it would take twice as long to train in a simulator as in real life.

When a trainee enters such a training system they have already proven that they have mastered basic flying skills in a real aircraft. This is analogous to the attaining of basic surgical skills. Importantly, certain 'what if?' scenarios can be enacted that could not be performed in a real-life training situation. The time taken to acquire surgical skills correlates poorly with the actual time allocated to learn such skills since only parts of operations may be learnt at each operation initially, and the cost of delays induced by training is considerable.

In order to provide a value for comparison between systems the concept of using an A-B comparison of the AIPES scores was introduced. This becomes more complicated as the tools only offer learning opportunities for the individual. The effectiveness of the simulation does not necessarily correspond to the effectiveness of the tool for generating a successful educational outcome. This is why the AIPES system extends the isolated simulation evaluation (AIP cube) so that the educational scenario and environment is brought into the system. This then forms part of the process for evaluation of the effectiveness of the educational tool compared with equivalent real world training opportunity exposure (e.g. attending an operating list) providing an A-B comparison – referred to as the 'reality index'. This is because the potential operating time may be less than 100% efficient and the educational tool may therefore provide an opportunity at least as good as real surgery for attaining certain specific discrete skills. Were the ratio of educational values of the simulated surgical experience compared to actual experience to be proven then it would be possible to refer to this as a transfer effectiveness ratio similar to that used in the vehicle training simulation domain.

If skills and experience can be gained that cannot otherwise be achieved through 'real life' training opportunities, then a reality index > 1 could be possible. The equivalent of this in simulation is to cut a nerve, (not intended in actual surgery!), or manage an engine fire or equivalent emergency whilst flying**.** It is still an arbitrary value and would need to be referenced to the TER to be meaningful. If this were done we would then have the chance to assess simulators with a small number of experts and reliably extrapolate the results to assist in cost-benefit predictions etc.

4.14.1 Relating the transfer effectiveness ratio to surgical simulation

The TER would measure a final outcome of a simulation (training) package. Aims for future evaluation include achieving a good correlation between the AIPES scoring system and the TER so that AIPES can be used to reliably predict the TER, allowing design phase modifications without having to run full clinical trials. It thus depends upon having an adequate number of people assess the system, which in the case of surgery requires a significant investment of time, money and human resource.

Simulation, at this stage of surgical simulation infrastructure development, is realistically only part of a broader training scheme (an adjunct) and hence the need to integrate it with the 'virtual university' flexible and distance learning infrastructure leads to the following list of aims for future evaluation:

- With the multiple **scenarios** that will need to be trained and the time required to train people in them. This points toward the need to develop a system to assess simulators that can be performed reliably and safely, with minimum effort and maximum reliability.
- To minimise the need for verification and validation upon training, the system should obviate the need for a full clinical trial each time a system is modified.
- The system needs to be acceptable to all parties so that it may be endorsed by official bodies, allowing for formal recognition of simulator training as part of the already established higher surgical training infrastructure – such as the Continuing Medical Education (CME) points system.
- Any scoring system to have predictive values must be related back to the TER.

4.14.2 Possibility of developing useful outcome measures

Usability testing of the user interface will lead to classes of simulations in a similar way to vehicle simulators. The perceived educational value will depend upon user uptake, and should have a role to play in protecting users from 'skill fade'.

4.15 Educational scenarios: learning issues

The demonstration virtual university training environment (VOEU) provides a rich source of material to enrich users' domain (declarative) knowledge in: basic science, orthopaedic disorders, and approach to patient, investigation and disease management. It focuses in detail on a vertical prototype (shoulder module) that provides a network of information that can be used by all readers in a browsing or goal-directed learning mode and is suitable for surgical trainees and practising surgeons. In surgery, skills-based competence is vital. It is the attainment of skills, not the duration of the training, which is important. These skills are:

a) Physical, e.g. equipment and procedures.

b) Cognitive, e.g. analysis, interpretation and decision-making.

It is the cognitive skills that demand a more sophisticated learning process if they are to develop (97) and with adaptive hypermedia mechanisms it is possible to match the content level with the student's knowledge through interactive case studies.

The problem-based learning (PBL) approach that is at the core of the VOEU integrated curriculum was adopted for VOEU as it can holistically represent the complexity of clinical reasoning. Trainees are then able to develop their procedural knowledge, which elaborates how surgeons approach problems, interpret clinical information and make decisions (98).

4.15.1 Developing cognitive skills through procedural knowledge

Some of these critical skills and processes that surgical trainees need to learn are decision-making, reasoning and problem solving (99): essentially skills that relate to a clinical diagnosis and management, which is a pivotal activity for all surgeons (100). To develop these skills, it is important for trainees to practise relating disparate pieces of information from a patient within the clinical context. This is the indication for simulation development. The graded case studies in VOEU provide clinical surgical education that develops trainees' clinical reasoning skills within the domain of Orthopaedics.

The *Master class* expert case studies were developed to encompass a wide range of information available in the clinical context. An expert practitioner generated these following a heuristic analysis of the clinical decision-making process. The evaluation of the user educational environment reflects the cognitive skills identified by Nkanginieme (100), namely that surgeons:

a) Obtain and recognize symptoms in the patient,
b) Identify the appropriate system involved,
c) Speculate on the pathological processes,
d) Differentiate pathological processes,
e) Identify the possible causes of the pathology,
f) Evaluate all pieces of information and make a clinical diagnosis.

To support trainee development of these cognitive skills, explicit procedural steps for the clinical cases can be presented: the referral letter, patient history, the examination, investigations, diagnosis and clinical management. On working through the latter stage, trainees are confronted with a wide range of options that should benefit from simulation support of procedural skill acquisition. This gives trainees information pertinent to the case under study. They make their diagnoses and then obtain subsequent feedback upon why that diagnosis is likely to be appropriate. This feedback, and the feedback across all case studies, makes explicit the expert's heuristics of clinical reasoning. This is an attempt to share that clinical expertise through a modelling or a cognitive apprenticeship (101) approach, and it displays the underlying principles and rationale of the clinical diagnosis and completes the process. This enables trainees to develop their cognitive skills, providing a framework for problem solving and hence clinical reasoning. The purpose of this is to bridge the declarative and procedural knowledge adaptively.

4.16 Conclusions regarding the value of AIPES

The overall aim of a simulated learning environment is to provide a rich source of information that effectively supports the learner through the adaptive features and allows users to engage with material more effectively. This is achieved through careful simulation (complex multidimensional hypermedia) authoring while maintaining a constant awareness of users' needs. The main strategies for developing an effective simulation were to:

- Provide simulation adapted to the specific surgical task.
- Design a simple practical system that provides for upgrading as technologies and databases improve.
- Increase maintainability of the application through the distributed architecture.
- Increase the usability and accessibility of information through simple human computer interface adoption.

Using the AIPES system comprehensive evaluation of the features was initiated to scrutinize the effectiveness of the simulation to achieve its goals. This evaluation process will give some insight as to what simulation-based learning in surgery actually benefits the trainees. Continuous review of future users and evaluation of their perception of the value of the various qualities of the surgical simulations will be supported by the client server architecture using a DRJ. This allows users to express their perspective managing the user administration issues for contributing to future trials.

The results of the usability trials of an integrated web-based surgical training tool represent a start upon a road leading to closure of the audit loop, where surgical simulators are assessed accurately with respect to whether they actually will deliver the educational goals for which they are intended. These have been subdivided to reflect the main areas of the trials, which consider the simulation and the environment in which it is situated. The ‘pre-study’ provided feedback for the design of the final study and its proposed methods for data collection. It emphasized the need for greater integration of end-users into the design process. It was therefore necessary to focus upon testing a web-based simulation package in light of the results of the pre-study detailing the testing of the educational infrastructure *i.e.*:

- Assessment methodology
- Internal structure and content
- Multimedia Educational Orthopaedic Module

All of this is based around an interactive course and the assessment methodology employed which considers the conduct assessment of trainee performance. The results demonstrate *the iterative design and development process has benefited from users’ participation in the development process*. By setting and attaining quantitative usability goals during the design process of EVW prototype 3, it was possible to engineer usability into products that may be derived from this research project as part of a rollout plan for technology implementation. The ease of testing using VOEU DRJ technologies recommends the system implicitly.

Testing versions with users early and continuously supports design iterations. By identifying users who are able to develop the system, running trials and analysing tasks, it is possible to set usability specifications for the developing prototypes of the simulation system. By testing prototypes a Usability Test Report is generated, which can be managed by the DRJ. It will therefore be possible in future developments to see if there have been significant improvements in design by comparing datasets. More importantly from the user’s perspective, it will be possible to distil the questionnaires to only use the questions that are truly discriminatory and so minimise the effort required for feedback and thus decrease the incidence of user fatigue with the analysis method. The difficult part is selecting and organising tasks to test, creating task scenarios that are realistic. The Zeltzer cube analysis partly demonstrates the future path of development based upon results so far, but as a system for analysis of integrated simulation it remains limited, and the proposed revision of this AIPES scoring for the EVW Prototype 3 has demonstrated a workable alternative.

Preparing other material for the test, as the AIPES scoring system demands, also depends upon preparing the test environment. It is necessary to study task analysis and quantitative usability goals so that timely issues based upon the earlier tests of this system answer the questions that remained. From the pilot study it was clear that a usability engineering approach and an iterative design approach was needed. This was borne out in the main studies. By developing user profiles, it is possible to select and organise tasks to test and create task scenarios (cases) that are suitable to the training grade of the trainee. The question is now how to progress from 'results' to 'use'.

The initial results argue the way forward, considering the process of integration into the conventional pre-digital surgical educational world, and the implications for surgery and surgeons. The move from embedded systems to integrated systems requires consideration as it has implications for surgical education. The key is how to ensure that the quality improves as a result of the implementation. Proof the simulation actually improves training would require cohorts of surgeons to potentially be subjected to inferior training and therefore it is developed as an adjunct to conventional training. Its benefit can however be demonstrated.

4.16.1 Metaphysics of VR for surgical simulation

As part of the introduction of VR to readers, it is necessary to define the boundaries, the potential limitations upon simulation, and the aims of simulation. This includes the potential limitations of the human ability to perceive the performance parameters by which they operate. The human performance aspects are detailed in Chapter 3. First it is necessary to discover what the user wants to perceive.

The traditional definition of virtual reality implies that the operator should be induced to perceive a risk of consequence associated with failure to achieve an acceptable standard of performance, where in fact no actual risk exists. The particular training environment in which the trainee is learning and demonstrating their skills can to some extent generate this. For example should the trainee be using such a simulator as part of the regular professionally supervised training schedule, then the validity of their performance would necessitate their attempting to achieve the best possible results. Anything less would be playing a game. Due to modelling limitations, scenarios will remain simpler than the clinical world exposes users to. There is no reason to believe that this poses an ultimate limitation, especially if networked systems are employed.

Ultimately one would hope that such simulations become the types of machines that pass an analogue of the Turing test. *By this I mean that it is possible to interact in such a way with the virtual environment that the end user would be unable to distinguish between the virtual environment and the real one.*

Environment: It may be important to consider other aspects of the simulation, such as a similar physical environment to the operating theatre. A draped mannequin is used to provide the shape of the shoulder, in order to produce some semblance of a physical simulation of the operating environment. The physical environment in which the simulation is conducted may contribute significantly to the scenario generated in certain situations, but exploring this is beyond the scope of this thesis, which focuses upon the approach to the visual display component of the simulation. Such principles were originally described in the early 1960s, the most notable example of which would be the "Sensorama simulator" (102).

Since the human mind is adept at "filling in the blanks", making up for the shortfalls in certain areas of the simulation, this attribute has been used to good effect by concentrating upon such interactivity as the ability to exercise control over the apparent direction of the vehicle (*e.g.* a motorcycle). By various means, the operator can have enough "studio/artificial reality markers" to be able to maintain the willing suspense of disbelief.

Clearly in an operating environment there are certain reality markers which people use, such as types of clothing to demarcate who is able to perform sterile tasks and thus able to work within the operating field, or (something which is not yet possible to take into account) the recognition of other people within the surgical team.

This last matter is something which is often overlooked but which in any real operating environment is a significant factor in the smooth running of a procedure, with regard to both safety and the satisfaction of all those who participate. Although this is not fundamental within the actual simulator itself, it has to be emphasised within the tutorial part of a training package in an attempt to achieve a degree of compromise.

Where the operator obtains a view of the operating environment indirectly via a monitor, the environment is already immersed. This is one area where terminology becomes a little hazy since the definition of immersion defines the relationship between experience and representation, in effect eliminating the syntax-semantics barrier.

Therefore such a system for shoulder arthroscopy need only be "semi-immersive" since the operator is in effect already immersed within the appropriate operating environment. There are of course other areas that need to be considered in this interface and these are detailed in the equipment section below.

4.17 Applying AIPES to different simulations

As to which approach is more successful, only time will tell. The digital technologies outlined in the previous chapters may offer a safe training environment for familiarisation with new surgical procedures. To assess this potential, a survey of user experience was conducted in July 1998 using the first prototype known as the Exeter Virtual Worlds (EVW) shoulder arthroscopy simulator. A pilot study was arranged. Access and control of the EVW virtual environment was attained via a 'java' based interface.

The Exeter Virtual Worlds (EVW) Shoulder Arthroscopy Simulator is designed to teach surgical pattern recognition skills via an interactive medium, using a thin client via distributed computer architecture. A Java-based interface was developed to enable the EVW system to be operated remotely from within a declarative (factual) web-based training environment. This was evaluated as part of the simulation system itself.

By using Java for the control interface it was possible for the surgeon to operate the virtual environment from a remote machine, using familiar windows-style controls which a focus group confirmed as adequate. They provided both subjective comments and objective results upon the performance of the simulation system and characteristics of the proposed user group.

The development of networking architectures was demonstrated to assist the assimilation of multiple sources and media. An appropriate interface is required to provide the opportunity to collect statistically significant data from such simulation systems. This might best be embedded in an educational environment, and this approach is therefore expanded in later versions, notably as part of the EVW Prototype 2. The early experiments allowed the following two questions to be explored.

1. To what extent can simulation adequately mimic aspects of surgery?
2. Is the scoring system a viable way of quantifying simulation?

The latter question, explored above (page 83), used the AIPES scoring system, so the pilot studies now concentrate upon the realism of the embedded simulation.

4.18 Pilot study methodology – nature of the trial

Since part of the iterative design process is to build new hypotheses from the results of previous experiments, analysis of the Pilot Study Results leads into the plan for the main study protocol development. To provide analysis of the early results for the preparation of the next round of experimentation, results were analysed according to this hypothetico-generating approach[21], the EVW Prototype 1 study was initially conducted to evaluate the potential of the EVW system prior to its integration into the Exeter Virtual Orthopaedic University (initial virtual university model). As a consequence of this work the hypothesis of this thesis developed, with the objective of further analysis in the form of usability assessment being to improve the system through finding out details about the user's preferences for:

- Visual Display
- User Interaction
- User Interface (UI) Design

To evaluate this it is necessary to provide Hypothesis-generating Descriptive Statistics. These Statistics will be descriptive only: and there will be no 'p-values'. There is no 'statistically significant' result. This may arise later in the hypothetical deductive phase of the work after this thesis. It is not necessary to have a fixed time frame for the recording of the trial data and so studies were performed at the convenience of the recruited trainees. Medical Research studies may be classified according to Table 4.2 below, which indicates the kind of data being interpreted. The nature of the pilot studies of the 1st prototype simulation is a descriptive, experimental non-randomised trial.

Descriptive		**Analytical**		**Experimental**	
Level 4	4	3	2	2	1
Prevalence	Incidence	Case control	Cohort, Follow-up	Non-Randomised Trials	Randomised Trials
Increasing difficulty as level of trial decreases					
Increasing cost as level of trial decreases					
Increasing strength of evidence as level of trial decreases					

Table 4.2 Medical studies design - evidence base levels

[21] The alternative to a hypothetico-generating approach is a hypothetico-deductive one, where the results are simply used to either prove or disprove the hypothesis.

4.18.1 Bias

This determines the minimum number of candidates that can be called upon to test such a system. Their lack of previous experience with surgical simulation and their familiarity with the facilities may introduce a negative bias as they may prefer what they find familiar. There are two potentially correctable sources of bias (Temporal and Geographical) discussed below that are to be addressed in the main study.

4.18.2 User profile - demographic information about the 1st prototype simulation (EVW) pilot study population

Because of the difficulty in obtaining background information upon the Information and Communication Technology (ICT) expertise and preferences of end-users this study was combined with a survey of the end-user population.

1. **Demography of the user group** All of the test candidates were practising orthopaedic surgeons in the South West Region of the United Kingdom. They are from the groups of Senior House Officers, Specialist Registrars and Consultants, and so are all medically qualified. The following demographic information was collected prior to evaluating the system. All candidates were male. This is not surprising, since at the time of the testing, only 25 of the 525 orthopaedic specialist trainees were female (4.7%). Orthopaedic experience ranged from 0 – 22 years in specialty with a mean of 4.3 years.

2. **Method of selection of participants** The study involved 24 trainee and expert orthopaedic surgeons from the South West region of the United Kingdom. It was held between the 10th and 30th of July 1998. The users evaluated were in the following clinical posts outlined in Table 4.3:

User Group	Number
Basic Surgical Trainees	6
Higher Surgical Trainees	14
Career Grade Surgeon	1
Consultant Surgeons	3

Table 4.3 Surgeons participating in the 1st prototype simulation study user clinical experience results

Participants were practising surgeons either ‘In training’, *i.e.* Senior House Officers and Specialist Registrars, or Consultants. All participants were volunteers. The selection process potentially biases through self-selection in that the candidates were invited to attend the demonstration session and then participate in the trial of their own volition. There was no pecuniary incentive for the candidates.

4.19 Discussion of the validation process

This section evaluates the results generated consequent upon using these evaluation measures. It prepares the reader for the synthesis of these indicators into the working model (103). The AIPES scoring system can be applied to the results of the 1st prototype simulation (EVW) Pilot Study.

This pilot study of 20 trainees gives us an AIPES score of 60% for the Exeter Virtual Worlds Shoulder Simulator. This is not an unrealistic starting figure for the performance of a first generation simulator: it does not fit well with the range of answers one would expect when one reads the comments made by those using the system.

"If the pictures were a bit smoother, and if there was somewhat more kinaesthetic (force feedback) input, it would be utterly realistic but as it is, it is already much better than a picture book about arthroscopy".

"I have no experience of shoulder arthroscopy but even after one session on the simulator, would feel much more confident in attempting a shoulder arthroscopy".

"I didn’t see major problems even with the jerkiness of the picture. Often it is difficult to get a good view of structures on arthroscopy, and a little movement can result in loss of image".

This study measured performance (using the AIPES formula) as 2 – 17% effective as compared to reality. This is referred to above as the *'reality index'*. In order to address this question, the main study will require the population to validate as well as verify during the evaluation.

4.20 Plans for surgical scenarios

The great advantage of computer simulations over real surgery is that surgical crises can be fabricated and the aptitude of the trainee surgeon to cope with these can be tested. This is the same as crisis testing in flight simulation: the ‘what if’ scenario, such as: ‘*What do you do if the wheels get stuck in the upright position in a Jumbo Jet?’* or ‘*What if the endoscopic clip slips off the sectioned cystic duct?’* It is said that aeroplanes are so automated these days that any given pilot will only have to think and act correctly once in his/her career. He/she is paid well, and assessed every six months in the simulator, so that when the decisive ten seconds of the pilot’s career arrive, he/she will act in the correct manner. Whilst surgery is not so automated, and indeed surgeons have to concentrate every second of their professional lives, crises may have similar life or death consequences, and it is vital that surgeons are trained to the same, if not better professional standards. Surgical simulation should therefore address not just basic skills but also evolve to teach the trainees how to manage adverse reactions.

Surgical simulation is still in its infancy. Progress is slow because there are several small and under-funded groups tackling the same problems independently throughout the UK, Europe and the USA. Satava's group (104-107) has suggested a timeframe for the adoption of widespread use of endoscopic simulators in the USA by 1999, realistic simulation of hand manipulation of organs by 2000, simulation of two handed ligation and suturing by 2002, adoption of simulators for certification by 2005 and full integration of simulators into surgical training and certification by 2010.

It is time that Europe caught up, and this will require action from the profession, the Royal Colleges and the National and European governments. This thesis is aimed at providing a blueprint for this process so that surgical scenario development becomes a clear ‘target’. This earlier prediction was based upon the principles of using video databases for the training and thus low interaction and autonomy were anticipated.

Factor		Shoulder
Autonomy	A	0.64
Autonomy Coefficient	a	"1"
Presence	P	0.64
Presence Coefficient	p	"1"
Interactivity	I	0.54
Interactivity Coefficient	i	"1"
Collective AIP Coefficient	c	"1"
Environment	E	0.60
Environment Coefficient	e	"1"
Scenario	S	0.56
Scenario Coefficient	s	"1"

Table 4.4 AIPES results for the 1st prototype simulation pilot study

Applying the AIPES equation with the denominator set to 1 as this is the first study and the effectiveness of the real world learning opportunity is considered uniformly standard to the pilot study results:

$$(20\ (1.\ A) + (1.I) + (1.P) + (1.E) + (1.S)) = \text{AIPES}\ \%$$

Equation 4.6 AIPES Equation

$$20 \times (0.64 + 0.54 + 0.64 + 0.60 + 0.56) = 59.6\%$$

Equation 4.7 Pilot Study AIPES Score

It must be emphasised that this system tests the simulator with respect to the tasks for which it was specifically designed. It does not test the simulator with respect to generic tasks, unless these are prepared as *'core'* scenarios. Such a philosophical issue is best resolved early in the development of surgical simulation. The surgical perspective is a pragmatic one: *'horses for courses'*. One needs to compare like with like, and as in surgery, there is no such thing as a generic system or scenario for operation, only generic skills, such as suturing or dissection. Strictly speaking, there never will be. This is unfortunate, since a benchmark would allow for realistic comparison. Well-described contrived constrained scenarios are likely to become the mainstay of training and testing. By developing a basic test system to compare across platforms, we can develop an assessment standard that is platform-independent. This approach was supported by the neurolinguistic programming analysis described in Appendix 7.1.

4.21 Conclusions drawn from the 1st prototype simulation trials

For a realistic extraction of knowledge from results it is necessary to view them in the social and political context, as mentioned in Chapter 1, following the Bristol Enquiry (34). Continuity of the arguments is maintained by the establishment of the following key points (conclusions) from the 1st prototype simulation and later trials, which are explored in the rest of the thesis. Thus the EVW prototype 3 trials needed to be performed to carry out the hypothetical-deductive analysis of the integrated environment.

This emphasises the value of the results as a marker of the subjective will for the further development of such integrated systems, focussing upon visual quality. Such results suggest that evidence-based judgement may lead to changes to clinical practice such as the introduction of relevant training tools into regular orthopaedic training courses. The arguments proposed in the author’s discussion are expanded in the Chapter 9. This will be supported by the results of the EVW Prototype 3 trial, detailed in Chapter 9.

The author considers that pursuing the concept of the integrated simulation devices is worthwhile upon the strength of the 1st prototype simulation and later trials and so the next stage of the work is expanded upon, in light of advances made through participation in the VOEU project and its derivatives, in Chapter 8. Three areas of contention remain:

- Inherent bias in a non-randomised study is addressed methodologically in later trials by providing support for trial scalability and asynchronous user data collection.
- Uncertain directional control of the hand held ‘virtual camera’ is eliminated by EVW prototype 3, leading to trials in the EVW Prototype 3 with a conventional 2D mouse for an input device.

- Value of developing a ‘flick book’ virtual environment is seen in light of the progressive improvement of graphics-based simulations. This remains a research issue for latest prototype. For the main trial the streamed bidirectionally linked video database was adopted.

Since the results are valid (page 103) and the hypothesis foundation justified (page 96) both internal and external validation exists. The limitations of the study are founded in the assertion that it is not possible to obtain statistically significant results from the pilot study (the 1st EVW prototype simulation trial). This raises the issue as to how one can get access to larger numbers of such test candidates. The conclusions are thus appropriate for the results if one remembers the context of the experiments.

4.21.1 Evaluation of surgical simulation components in the 1st prototype simulation

Surgical simulation is still early in its development. Without a rational integrated policy, it is likely to be under-resourced and under-utilized. A coherent research and development strategy has been proposed which advocates bringing technologies and expertise from the different disciplines together, so that new systems may be exploited for the benefit of the public. New surgical tools and techniques will increase the demands to achieve cost-effective training of surgeons in a safe environment. Simple and reliable evaluation is key to the integration of these new resources into the established training infrastructure. Underpinning this is the need for an accurate virtual training environment, specifically, its integration into the tutorial systems used to train surgeons.

4.22 Summary of the 1st EVW prototype simulation work

The EVW Prototype 3 study protocol development was taken forward to the VOEU project based upon these early results, and is described in Chapter 8. Simulation can adequately mimic certain aspects of surgery, borne out in both subjective and objective information collated from the pilot study. Also, the AIPES scoring system offers a viable way of quantifying surgical simulation, at least with respect to the concept of minimal access surgical simulation.

A coherent strategy for the development of the surgical training infrastructure is vital. A concerted effort was made to develop and uphold an ethical code, detailed below in Section 9.10. Over-regulation strangling the potential new technologies, and restricting both their use and their development, is to be discouraged. This is a fine line to walk, but should be achieved through the preparation of minimum standards for the accuracy and performance of tools. The VOEU Project should do just this by allowing the continued development of simulation systems, independent of each other, but with access to a shared graphical user interface and working environment.

4.22.1 Surgical simulation should not be seen in isolation.

The changing educational context, such as proposals for modernising medical careers, poses problems with the concurrent development of other tools and techniques. This affects both Computer Assisted Learning and Surgical Practice. This will necessitate reviewing how learning environments and simulation systems may be integrated to complement each other. Virtual environments used for training will need to meet a standard that both provide an adequate training environment and also an adequate testing environment.

Methods of communicating this resource to those who need it is a major part of this work, since this responsibility should not lie at the feet of individual patrons or trusts, but should be accommodated as a collective responsibility.

To use the analogy of the flight simulators, it may be commercially the responsibility of the airline companies, but the Aviation Authorities set the standards to which pilots should be trained and define the acceptable standards for the training equipment. Such guidelines should be established for systems that are going into clinical training environments.

4.9.1 Unique properties of VEs can create novel usability problems

The number of users testing the simulation resources is limited. In the case of the 1st EVW prototype simulation study, this is based upon the availability of volunteers from one region in the UK. Prior to the main study (the version number 3) it is necessary to power the study to ensure that an adequate sample size is arranged to determine statistical significance. Using clustering, the sample size increases rather than decreases, although this will be necessary to consider the three basic groups – Novices, Trainees and Experts.

The sample sizes for the pilot studies are usually small and so do not reflect a stratified population well. If one is going to develop the system fairly, it will be necessary to compare it with other systems on a like-for-like basis, and probably this means comparison of performance with the following demographic variables:

- Grade (indicative of stage in surgical training)

- Year Qualified (Relating to years of clinical experience)
- Years in Specialty (Relating to years of clinical experience related to the specific subject)
- Number of Shoulder arthroscopies Previously performed (Relating to specific sub-skills experience)
- Number of Minimal Access Surgical Procedures previously performed (Relating to specific sub-skills experience but unfortunately not recorded in this study)

Even stratification into the three groups mentioned above may not be accurate. In effect there are two distributions that need comparison, the general ability level of the individuals:

- The performance of the average surgeon matched to the population of 'average surgeons'.
- The performance of the simulator - acting as both a teaching and a testing tool.

Resolution of the resulting paradox (see Section 3.8.1) is the principal indication for the main study, though in performing the study a conclusive result is by no means assured since the power of the study is not known, and the tests may be poor discriminators. VOEU offers the opportunity to assimilate the data to evaluate it long term. The following problems arise:

- There is no previous dataset with which to compare.
- Changing training scenarios will introduce other variables.
- Valid outcome measures are needed for endorsement of new results.
- There is a need to embrace the conventional training infrastructure.

If it is possible to demonstrate the potential of such training aids as reliable tools for evaluating surgical performance, the field opens up a global market for the use of virtual environments within the already established training infrastructures. It does however continue to involve users in the design process and also imposes both broadband network requirements and agreements upon preferred network protocols.

Virtual environments used for training will need to meet a standard that provides both an adequate training environment and also an adequate testing environment as users become familiar with systems and can contribute material themselves.

4.23 Summary

Simulation cannot exist in isolated splendour. It must be seen as part of the integrated educational infrastructure. By bringing the rest of the infrastructure and the individual up to a functional level they will be able to benefit from such integration. This has been every successful for companies in the aviation industry such as British Airways. Their training system can be used to train pilots to co-pilot specific aircraft without ever having flown real aircraft of that type. There is no reason to believe that the same could not be achieved in surgery. A simulator is only going to be as good as the design specification that defines it. Accuracy of virtual environment generation is critical. The specification must outline the aims and objectives for which the simulator is designed, and this thesis demonstrates the pedagogy into which simulation will be integrated. It is widely accepted that simulations will evolve through iterative design, taking into consideration demands and technological progress (seemingly ever improving components) pushing forward the ‘cost-performance’ envelope.

Before developing a new simulation system, evidence of the intended validation and verification procedures is required to ensure that the simulator will indeed match its intended goals. These will include the establishment of credible standards involving the selection of the test methods, and addressing the methods for their verification and scoring. For each simulator it will be necessary to examine the purpose for which it is intended, and thus define what is to be tested. The rationale for surgical skills assessment includes the summative assessment of these specific tasks. It is ultimately intended to help the development of tools for the maintenance of competence, as with the aviation industry. These will find a niche alongside the more conventional tools for the teaching of surgery, including plastic models, plastinated anatomical specimens, and illustrative art.

Equipment is only one part of the equation and so attention must be given to the development of the environment in which it is to be used – its educational and physical context. This draws upon the problem-based learning model for medical education that suggests the integration of such simulations into scenarios will produce a better resultant training tool for the surgical trainees. In the EVW prototype 3 simulation design a greater shift toward portability of the media to meet perceived requirements had consequences in other ways, such as trading access for interactivity, and the significance of this is in adopting the pragmatic approach opening up the field of pervasive simulation. The AIPES system provides an instrument to evaluate progress towards this.

5 *Surgical and educational factors for a distributed learning environment in orthopaedics*

5.1 Introduction

It is likely that peer review will remain the most reliable indicator of performance during training for the foreseeable future. The pilot study does not seek to undermine this, rather to complement it by providing a useful tool, developed to support the development of an individual's performance, in conjunction with the development of multimedia tools to guide the trainee.

From the early evaluation of the interface relying upon the two platform-independent systems, the role that Java can play in system integration will be 'outsourced' to COTS programmes. This eclectic tutorial-based approach will allow for different simulation systems, each with capabilities to provide training and evaluation in specific sub-skills. They will need to contribute data for evaluation. This will facilitate statistical analysis of a suitably-sized population of surgeons in later trials. The data of simulator performance will progressively improve in accuracy and ultimately provide a system of assessment with the underlying foundation of a large population of surgeons, due to the use of distributed computing architecture to develop this important database. The 1st EVW prototype simulation demonstrated that Java is versatile, though with limitations, the major advantage being that the designers are not tied to a single computer architecture or operating system. As adaptive hypermedia (AH) has developed, this philosophy has extended beyond a single language.

The results of this EVW Pilot study suggest that such interfaces are to be built into educational environments to assist the control of training simulators. They include the development of the *Virtual Orthopaedic European University (VOEU)* (96) and the continued development of the shoulder simulation system into the next prototype. Ultimately such controls can be developed into standards in their own right. The aim of future work is to ensure that these devices assure compatibility between the intra-operative tools developed for the training and operating environments and both objective skills tests and the subjective comments of the trainees are a promising start for the integrated system concept. For this a distributed learning environment is required.

The pedagogy relies upon an ontology developed collectively. The intention is for end users to engage appropriate material presented using the most appropriate media founded upon a user-centric model. Focusing upon computer-based learning:

- The feedback and tutor support provides a foundation for peer review
- This includes students, tutors and other consultants
- The ontology is pedagogy driven, and is outlined below.
- It is based around adaptive hypermedia

The map in Figure 5.1 provides orientation as one navigates the components in the integrated VOEU curriculum. This was constructed in various work packages (WPs) and so relies upon interaction of the different components. These may be from different sources and thus use a distributed architecture.

5.2 Surgical factors

The evolution of the medical networking infrastructure will dictate the rate of progress. At present the user could expect to have broadband pervasive access within 3 years and full integration with electronic patient records by 2009. The DOH is proposing over £32 billion funding for NHS ICT over the next decade.

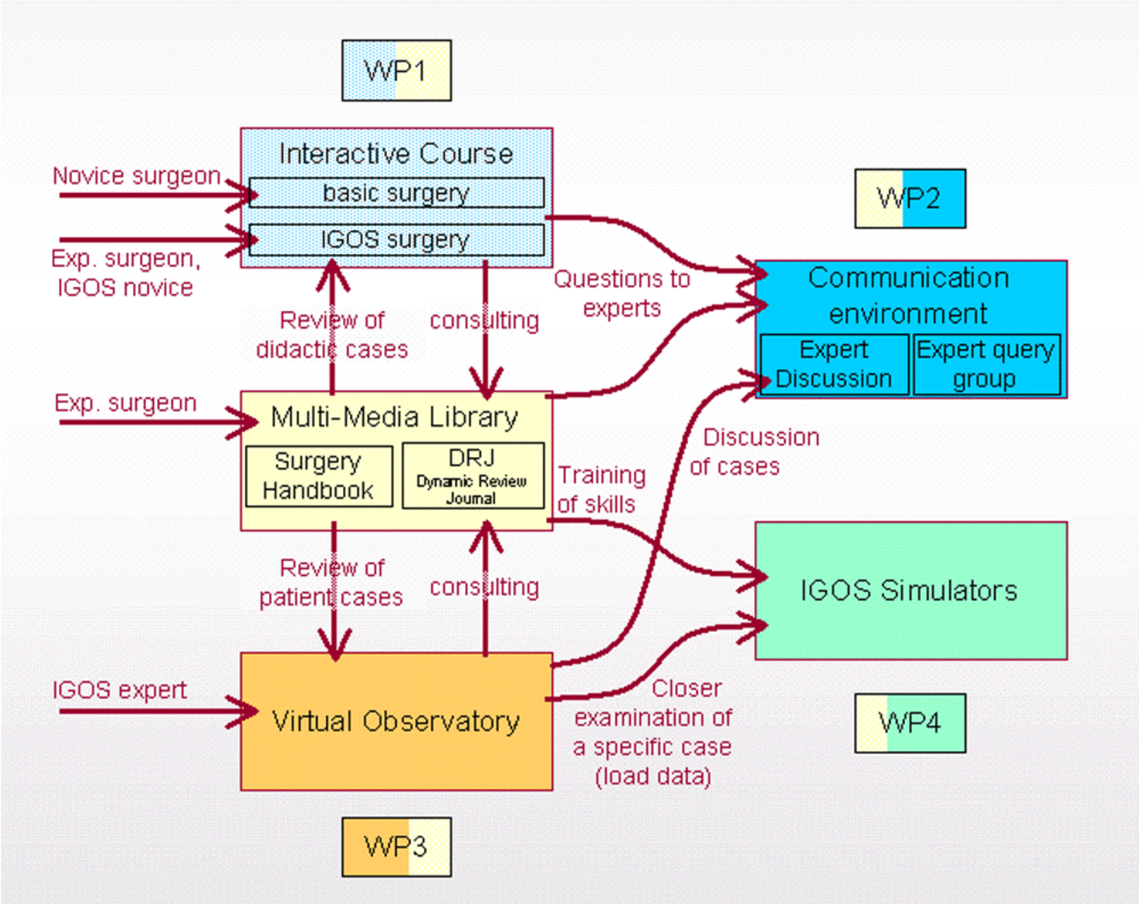


Figure 5.1 The components in the integrated VOEU curriculum

5.2.1 *The orthopaedic syllabus*

This is arranged by sub-specialty. Version 1.1 of the syllabus was a static entity. It had no relationship to the learner's previous experience beyond tacit acknowledgement of the assertion that the trainees must have attained entry-level knowledge to attend the courses. The learning agreements therefore allow the trainee and the tutor to agree the plan for the clinical experience and select suitable posts of a level of experience for the trainee to use for structured training purposes. The individual contract must be specific for the clinical post and the template can be used only as a recommendation. It is not prescriptive. The XML (108) foundation presents the user with data in a form that is interpretable for their level of knowledge and experience. The system is fully scalable with respect to the number of collectable cases. It records a permanent archive of clinical procedures. Because of the sensitive nature of the material, the security issues are addressed at the server level and consent must be obtained from the patients for their case material to be included.

The construction of metadata standards for the core components that abide by already approved standards such as the Dublin core (109;110), Learning Object Metadata and Information Management System (IMS), leads to a philosophy where components are 'living entities' whose survival within the university infrastructure relates to their relevance to users. Applicability and expandability (68;111) are vital characteristics of components in VOEU. This allows for educational presentation to change and the style to be adopted seamlessly. The same is true for the access to different views of datasets used in clinical trials. There are three underlying motivators for the *Information Society* philosophy being applied to orthopaedics:

- **Constantly updating knowledge base** for both procedural and declarative learning. It is said that to remain truly abreast of the developments in the field of orthopaedics a surgeon would need to read approximately 1000 technical papers per month.
- **Limited user time and computing expertise** as demonstrated in *the 1st prototype simulation* and to a lesser extent in *the 3rd prototype simulation*.
- **Specialist knowledge base** varies according to user experience and its application.

Based upon work by Prof. Bulstrode[22], it is necessary to classify clinical knowledge into:

- **Essential:** Part of the core curriculum relating to safety issues. Every trainee must be evaluated and demonstrate passable skills on every occasion. Should the trainee not be able to achieve a satisfactory standard then they will not be able to be awarded the relevant qualification. There is no negotiation with respect to these.
- **Important:** This is also part of the core curriculum. Although time is not formally available to test this, it will be included within the questions of the core modules. The varying ability to pass this will constitute the grade of performance.
- **Further Reading:** There is the need for a course to contain material that demonstrates various issues around the subject. In particular it adds background to assist with reflective learning. This does not need to be evaluated formally.

[22] Personal communication via 'training the trainers' course.

5.2.2 *Sources of data for educational purposes in VOEU*

This background, blended with the variety of user experience encountered, requires a reliable and robust system for the collection of data from both the real-life and simulated environments. The data will then be used for both teaching of novice surgeons and the comparison of effective training techniques for different types of surgical intervention. The role of the user profile and its implementation are important. This is relevant to the user as the content is adapted according to the personal profile. The approach adopted is to collect this data from four main fields. These are:

1. *Multimedia Educational Modules*, which provide the declarative (factual) base of material for the education of the users.
2. A *Virtual Classroom* environment for exchange of views, and monitoring of progress.
3. A *Virtual Observatory* for the collection of data from simulation systems and the actual intra-operative data collection discussed as the DRJ.
4. Novel Modalities of *Simulation* (36;112) for the emulation of surgical procedures for training and experimentation focusing upon micro-surgery, exemplified here by the shoulder surgery simulator but with other teams developing hip, knee and spine simulations.

By combining the above disciplines within one working educational environment, the virtual university infrastructure (113) aims to meet the needs of clinicians combining clinical, educational and research duties. VOEU was specifically established to address the issues of data access, presentation and development of a structured learning environment for the training of novice surgeons. The technologies that support such a working environment for familiarisation with new surgical procedures and the management of clinical case audit can be made available to all stakeholders for lifelong learning. This contributes to the VOEU philosophy of embracing all possible learning models by providing learning object metadata, which allows course convenors to build their own course structures, whilst focusing upon the problem-based learning model (89;114) to allow multiple modalities to be presented to the trainee.

Trainees will be shadowed by this digital record accrued over time, formalised as the surgical education ontology based upon the HST Infrastructure (Figure 5.3), of which a part is the set of learning agreements outlined above. There is a core of documents and components linked into the VOEU surgical educational learning agreement. This will continue to be refined as part of the VOEU pedagogy.

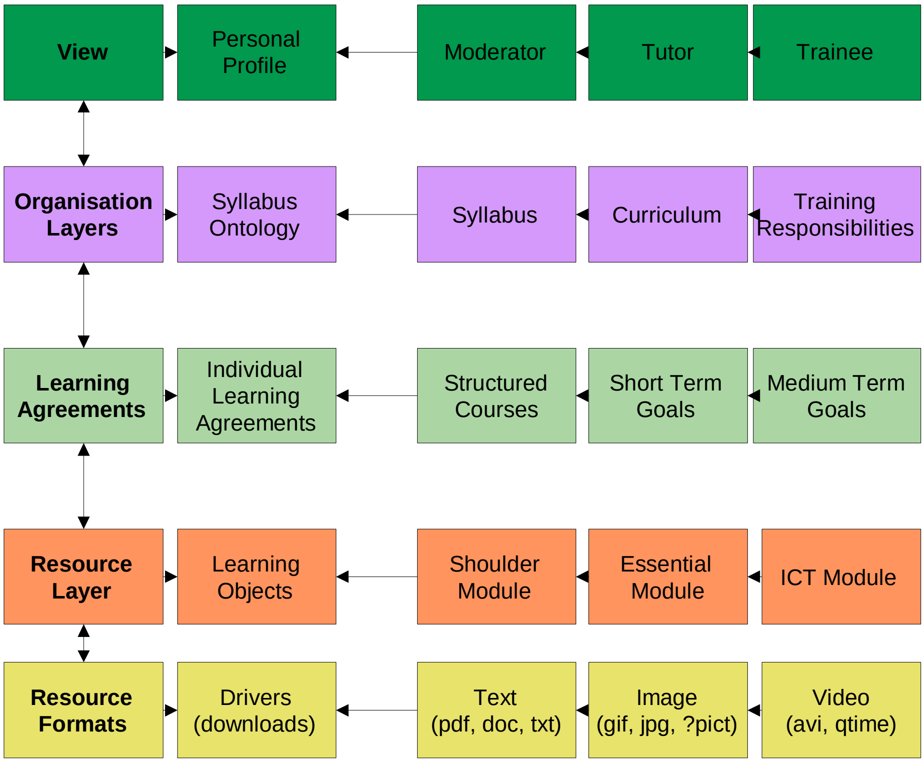


Figure 5.3 Electronic higher surgical training infrastructure

5.3 The 1st prototype simulation (Exeter Virtual Worlds) surgical simulator shoulder arthroscopy tutorial

As stated in the introduction surgical skills are learned by the progressive honing of sets of skills creating *'islands'* of knowledge. The experience gained using a specific simulation represents one such island. It considers both the training material and how this is integrated into an interactive multimedia educational orthopaedic module. The software application must provide a user-centred perspective upon the environment.

5.3.1 The use of Java to improve the interface

Provision of a web viewer-based interface created an upgradeable platform-independent UI that will minimise the risk of rapid obsolescence of applications. As the new formats evolve more rapidly than the standards that may unify them, the commercial companies that are market leaders, are driven to develop platform dependence as a way of protecting their investment and developing brand loyalty.

The need to confront this upon the grounds of unbiased clinical practice supports Java and other standardised packages. It was therefore decided by the consortium of VOEU to adopt the W3C standards for *eXtensible Mark-up Language (XML)* in order to address this issue.

5.4 Development of the VOEU visual integrator

XML offers the option for handling different media in different languages that cannot be done easily in other ways. The users can customise the system by allowing for different natural languages to be integrated into the databases that are accessed. The reason for extending effort to build a specific *Document Type Definition (DTD),* or schema that then becomes public domain is that the system allows for the following three major issues to be addressed:

1. ***Extensible*** *mark-up language* is just that. The current boundaries are not fixed, it does not know what structure will be required in the future, and need the flexibility to update and modify the structure of the language accordingly. This will make the system more robust in the future as it is able to accommodate changes in other disciplines, and in orthopaedics itself, allowing for such features as user customisation, interactive menus and scenario access. The system will more readily accommodate different natural languages and provide adaptation to different hardware specifications.
2. *Extensible* ***mark-up*** *language:* This is a recognised standard for the evolution of data structures, which led originally to the structured generalised mark-up language (SGML) and thus to HyperText Mark-up Language, HTML. This means that much of the basic work of HTML is readily applied and where possible, by adopting standard principles used in other healthcare application areas such as HL7. It is possible to assure uptake and evolution of the standard in harmony with other abutting fields of research and healthcare delivery development.
3. *Extensible mark-up* ***language****:* The basic framework of the XML language satisfies the need for platform independence. The future development of Java may support this when one considers the different media options also available. XML accommodates the following constraints. It is necessary to place this in the context of two axes:

- *Java spectrum – types of devices* -The development of the spectrum of devices that can process and deliver data to users provides a User Interface with a uniform application environment for healthcare information delivery. This is part of the interface for the DRJ.
- *Limitations of Document Type Definition potential range of applications* - This work is not the remit of this thesis but some of the principles were developed in association with the VOEU project. The domain under the influence of this project was a small subset of that which is potentially possible, leaving the option of developing further modules and schemas. Ultimately these should be configurable using a form (*.aspx*) input system, currently proposed (115).

5.5 Virtual orthopaedic university overview

The evolution of the information society technologies is having a profound impact upon the field of Orthopaedics. Recognizing the need to assist the life-long learning requirements of this field, the European Commission has funded the VOEU Project. This affects all aspects of Orthopaedics using computers for information management. As a university, its services will support education, research, administration, and clinical management.

To provide a framework for the management of the multiple media formats including documents, which convey information, and the need to clarify data in the virtual environment, a version of the *eXtensible Mark-up Language (XML)* dedicated to Orthopaedics was developed. This XML proto-standard was originally referred to as *Orthopaedic Surgical Mark-up Language (OSML).* Although earlier work (116) introduced the structure of OSML, including its advantages over standard *HyperText Mark-up Language (HTML),* and issues related to appropriate ratification of standards, the author perceives this as an intermediate step in an evolutionary process, with later versions including both generic and specific schemas.[23]

OSML may form part of a wider ***Generic Surgical Mark-up Language (GSML)*** which itself could form part of a general ***Health Care Mark-up Language (HCML).*** The structure of OSML must therefore accommodate the potential for a GSML based upon an ISO standard (117). This implies that the OSML standard should be consistent

[23] HL7 (www.hl7.org) consortium in the United States of America has also attempted this. Endeavours to agree standard document ontologies prolong the process of evolution.

with further specialist surgical mark-up languages as they are developed. Thus the user could access material not specifically developed under OSML, if it is relevant to their work in orthopaedics. The OSML specification should further support the automatic placement of files with the defined file structure (108) using conventions currently used for standard surgical reports. Examples of current developments can be viewed at the VOEU website (http://voeu.ecs.soton.ac.uk[24]). This is still in an implementation phase.

The design of the VOEU file structure views the route of access as starting with the top level of the proposed HCML. In certain areas HCML will expand to provide GSML, which will then offer the opportunity to cover the areas that are used in all surgery. There will be special attention directed toward orthopaedic surgery using OSML, which is the focus of this work. The likelihood is that as other efforts to organise and ratify a generic surgical mark-up language are initiated, the OSML will act as a compatible basis since it adheres to the XML principles, and thus can either stand-alone or integrate with GSML and HCML script.

The present way for achieving this is only via search facilities, and according to the GRID philosophy (118) this needs to be extended beyond keyword searching that is currently employed. An approach to this is the integration of the ***structured surgical course model (SSCM)*** structure with the hierarchy of the keywords used for indexing, being referenced to the syllabus (as derived from the BOA core educational content for Higher Surgical Trainees in Orthopaedics and Trauma document) to build a context-dependent search system.

5.5.1 The 'Visual Integration' of information presented to the user

To develop a co-ordinated tool for the Exeter Virtual Orthopaedic University (initial virtual university model) User Interface, it is vital that off-the-shelf components are seamlessly matched with components developed by the in-house design teams. This process requires the acceptance of standardisation of various rules from the initiation of the system. The University of Southampton (the 3rd EVW prototype simulation) system is integrated with multiple other components developed by multiple sites. It therefore is imperative that the approach to integration is separated from the resources - a middleware issue. By adopting style sheets (*.xlst)* it is possible to ensure that all groups

[24] I acknowledge the discussions of this issue with T Bunker, G. Jones and N. Campbell, which led to the use of the structured surgical course model development.

may develop their own end user applications and allow them to be viewed by the user through a common interface.

If a user is to instantiate a new version of the interface each time they log on, then the institutional look and feel and user profile need to be available for their use. The author therefore proposes the development of an *'Orthopaedic Outlook'* – similar to the popular '*Microsoft Outlook™*' integrated visual working environment. This needs to accommodate everyday administrative functions, and also special course-related features. The Helmholz Institute has since developed this system further in the VOEU as the Visual Integrator (119). The following list expands this concept:

- **Institutional banner and buttons –** identifying the licence holder.
- **User profile –** The user requirements must be met for this, the individual's preferences, passwords, authorisations etc, with an option to edit. This should ultimately include links to other sites that users have rights and privileges at. An example would be the use of archives for journal articles held upon different servers. This includes the tools that can later be developed into part of a personal agent cloud for the development of resources such as a personal 'Dynamic C.V.' data assimilation service. Also included would be an educational schedule and Examination records. The user interface must be able to launch the relevant applications promptly, from an appropriate database.

5.6 OSML development

The plan for the development of the Multimedia Educational Orthopaedic Modules set the milestones of OSML DTD, OSML documents, OSML editors and an OSML browser/viewer. After initial definitions of the Document types, and use of *xmlspy*™ to edit these, the consortium adopted the path of developing schemas for wider compliance. Let us first consider why we are adopting an XML development approach was adopted.

5.6.1 Potential beneficiaries of OSML

There are several categories of users of OSML: they will use dedicated editors to manage documents and other media. All are likely to be primarily medically trained and so the development of the editor should provide a 'what you see is what you get' (WYSIWYG) working environment. This is achieved in the VOEU demonstrator by using Microsoft Word™ as the editor and then uploading to an XML converter with protection of the metadata in the form of a word template table for Dublin Core data collection. The content providers and users are presented in Table 5.1 of stakeholders below:

Stakeholder Group	**Category**	**Abbreviation**
Content Users (Learners): users of OSML repository	Medical Students Basic surgical trainees Higher Surgical Trainees Continuing Medical Education / Continuing Professional Development Other medical professionals[25]	UG=Undergraduates BST HST CME/CPD GPs
Content Providers: users of OSML editors	Surgical knowledge engineers Surgical curriculum developers Surgical researchers Medical researchers Administrative staff	

Table 5.1 Stakeholders of the virtual university

5.6.2 Off the shelf component library

Where possible, the aim is to use software *e.g.* MS Word™ generated .rtf files and human 'agents' from already existing resources, for the following reasons:

- Leverage in development time and refinement
- Familiar 'intuitive' best-of-breed design for the main components
- Open-platform design where possible
- Cost-effective licence distribution
- 'Automatic' adoption of the standards developed as part of XML (www.xml.org) and HL7 (www.hl7.org).

To develop the work of the formation of supporting multimedia-based educational orthopaedic modules that are managed through the visual integrator, it is necessary to define the tool components required for the delivery of the planned media and the collection of appropriate feedback data. These are categorised in the list below:

- Declarative Knowledge Content (XML documents, repository)

[25] e.g. General Practitioners. Operating theatre nurses etc., may be included as the project expands.

- Orthopaedic Surgical Mark-up Language (DTD)
- OSML editors/browsers/viewers
- XML processors/parsers
- XSL editors
- Link management tools
- DTD editors / authoring / modelling tools

5.6.3 OSML as a child of HCML

The concept of developing OSML as a schema calls upon the HCML (HL7 derived) schema. This would mean that we are borrowing from a major body of development and those components that are relevant to our project, whilst ensuring compatibility with the appropriate standard where there is overlap (Figure 5.4 and Table 5.2). An example of this would be the proposals by HL7 for the handling of DICOM images.

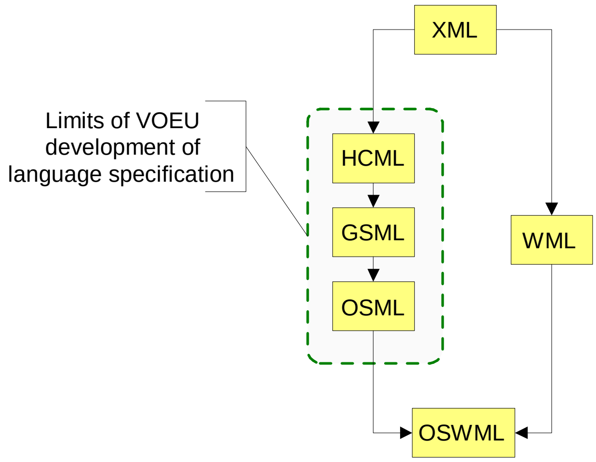


Figure 5.4 VOEU responsibilities for XML derived language development

Key:	
XML	Extensible Mark-up Language
HCML	Healthcare Mark-up Language
GSML	Generic Surgical Mark-up Language
OSML	Orthopaedic Surgical Mark-up Language
WML	Wireless Mark-up Language
OSWML	Orthopaedic Surgical Wireless Mark-up Language

Table5.2 Subsets of XML derived mark-up language

5.7 Designing for the desktop overview

The VOEU project provides the service over the desktop. This is the current preferred mode of data access due to the cost/performance characteristics. The paradigm shift is however toward wireless handheld devices as the performance increases, for the convenience of mobility and thus pervasive access.

Designing for the future means that the option to ensure compatibility with wireless systems is considered but it is not possible to develop an application interface for the mobile devices within the framework for the VOEU project, though the design has the potential to do so in the future.

Browser components: Let us therefore consider the required components for the OSML Viewer/Browser, which are a server script converting XML content to HTML, an OSML plug-in for conventional browsers and generic XML browsers when available or a dedicated OSML browser. The OSML DTD/schema development process focused upon three stages with an iterative design:

- Rapid prototyping using a story-boarding technique: pen and paper, post-it notes and a wall
- Drafts revised by the users employing an adaptive archiving system with versioning.
- This will lead to a draft ultimately becoming available in the public domain as an ISO standard through public identifier issued by the American Graphic Communication Association. This is still to be completed.

Content: The development of Structured Surgical Course Model (SSCM) documents/repositories is ongoing. However, the prototype is in service. Whilst the test material can deviate from this in accordance with the author’s wishes, the principle of having a template for the content results from co-operation with the knowledge engineers (surgeons and engineers). Content providers can use OSML editors to construct repositories. In the early the EVW Prototype 3 work an MS Word™ template supporting the metadata was adopted. Those with a surgical specialist background, and surgical curriculum developers are accommodated. The documents will be valid XML documents (OSML + other relevant schemas) and once vetted by peer review they are transparent to the end-user (though not strictly WYSIWYG…).

Additional file formats to be supported: These will need to be accommodated when demand arises and so should be formally recognised or specialised:

- ***Personnel Files:*** Personal Profile with security clearances (.pps), Mail files (.eml), personal contributions *e.g.* Archived Presentations (.ppt)
- ***Logs:*** Videoconference logs, Surgical Logs (.xls)
- ***Simulation specific files:*** Simulator Database (.ogl), Haptic Rendering (.ffb)

5.8 User interface

A Word *.doc* to *.xml* converter for document uploading (IAM Group in Southampton) was adopted. Preparation of the language required co-operation between end-users, engineers and educationalists, adopting a user-centred evolutionary approach with rapid prototyping and with regular evaluation including early user involvement. A range of methods for functional, data and usability requirements specification, including questionnaires, interviews and walk-throughs (paper prototyping) were employed. Other methods (not employed here) include heuristic evaluation, empirical usability testing (constructing scenarios with think aloud technique), monitoring (logs) and survey type feedback forms. Since it is recognised that adding another layer can solve the problems, the following structure in Figure 5.5 was developed:

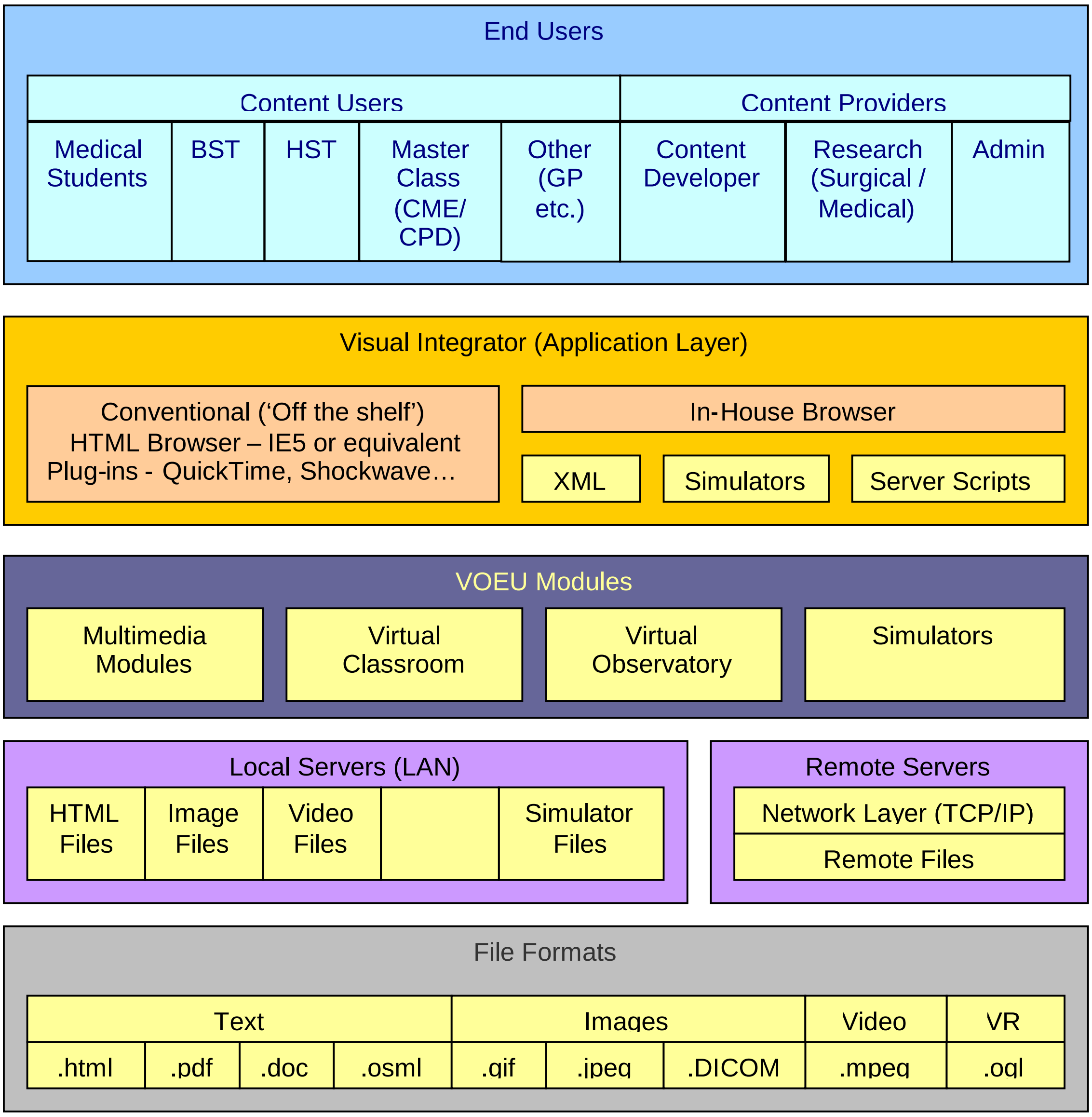


Figure 5.5 VOEU structures for the visual integrator

5.9 Orthopaedic Surgical Mark-up Language development

The file formats shown in Figure 5.5 should facilitate the advances in linking documents and pointing to components within them. OSML style sheets will produce a page layout that allows strong internationalisation capabilities. The 'Namespace in XML' recommendations published by the World Wide Web Committee (108) (W3C) have been adopted. This includes the capability for the file structure to be moved easily between servers without requiring editing of inter-document hypertext links. This is particularly important in the context of enterprises such as VOEU, which will necessitate the use of mirror sites, some of which will need to represent the data in different languages.

OSML will allow the integration of tools for searching databases that may be incorporated into all levels of training environments. The authors perceive the adoption of this language will assist the seamless integration of information technologies into the operating theatre environment and, as with all XML applications, OSML will be excellent for adaptation of the teaching material for handheld devices, hence increasing accessibility.

Early work on the *Exeter Virtual Orthopaedic University* (113) based upon HTML has demonstrated the potential for Java-based integration of the declarative learning material for either Intranet or Internet distribution, and the control of local simulation systems for surgical training using such a system. The future development of OSML will incorporate these underlying principles, though it will still require the activation of dedicated scripts to interact with the simulation systems (Figure 5.6).

Conclusions drawn from the development of OSML: The OSML provides a flexible language upon which to base media archives and communication for the evolving virtual university. The 'bottom line' is a robust document-handling process that supports peer-reviewed secure and appropriate placement of documents. The consortium of partners ensured that the following milestones were met by achieving the deliverables expanded below:

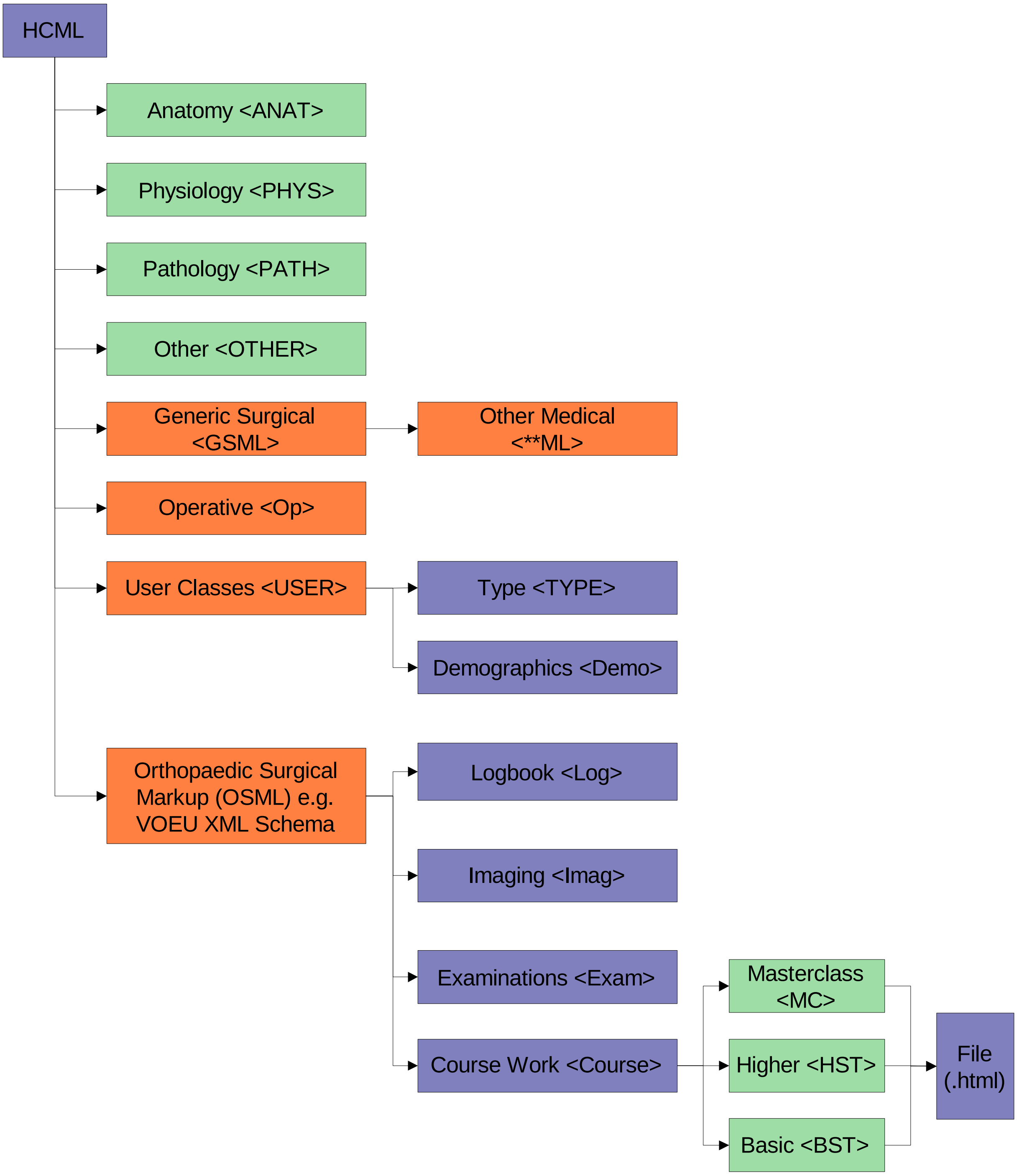


Figure 5.6 Using a schema to link mark up languages by .xml to .html resources that can be presented to different users (Colours assist in highlighting the hierarchy, with no specific significance)

1. **OSML DTDs.** DTDs were adopted for a working model of document development (These were later converted to Schemas).

2. **OSML documents.** Inclusion of the standard file formats into the DTD/schema ensures that the surgical knowledge base is accessible via the Visual Integrator. In the later versions, a schema writer was added that allows users to prepare more trials using variations upon the schemas but accommodating all the variables the users wish to employ. OSML file extensions will allow for the tailoring of specific courses to meet the needs of individuals, which are expanded upon in the pedagogy section of chapter 9.

3. **OSML editor and uploading process.** The aim of this was to progressively enhance the collection and preparation of educational material by knowledge workers (particularly Surgical Specialists). The process of automation of OSML Tag generation should facilitate the preparation of future modules. In the case of the developed uploading process, a metadata header was added to an MS Word™ document so as not to alienate the medical profession who, as seen in Chapter 8, are familiar with this software package. The editor will ultimately become part of the fully integrated toolkit for users, probably managed entirely as a web service, but this is at present solely a dissemination and exploitation plan. Richard Lawley (SOTON) built the current (working) system.

4. **OSML browser/viewer**. For the prototype of the VOEU Visual Integrator, the IAM group Virtual University systems adopt the current generation HTML browser (*Microsoft™ Internet Explorer v5.5* or above). Browser plug-ins have been used to provide appropriate functionality for specific file types. The aim ultimately is to be able to provide a bundled package that is downloadable so that end-users can automatically upload a suitable browser plug-in kit for the development of future Virtual University content. The planned features of OSML include:

- Multiple Applications for Multimedia components – this should release the educational content from the delivery media.
- Integration of Multilingual Systems – an enhancement of the visual integrator.
- Narrowcasting potential for dedicated operator skill levels in surgery and computing
- Reduced centralized server storage and data management costs via a distributed architecture
- Ease of operator upgrading with automatic link maintenance

- Ability to handle declarative and procedural material through ease of database access
- Surgical Log cross-checking for error correction
- Authentication and peer review ratification
- Integration of surgical and medical coding for clinical management – conforming to the proposed HL7 standards where possible.

The disadvantage of such a system may lie in the need to ensure acceptance of the specification that will necessitate the formal review process of the ISO. This may delay the implementation initially, but is ultimately a vital step in the maturation of OSML v1.0 from *de facto* standard to a ratified one.

5.10 Summary

This chapter addressed the issues of accommodating surgical and educational factors in a tutorial system for surgical training. The feasibility aspects are more important than the operational aspects here. Reference to standard models for User Interface (UI) design includes the use of ‘frames and style sheets’. It includes analysis of user-centred design and its significance to providing interfaces dedicated to surgical users. The interfaces dictate Staff Training Protocols, so that setting up new systems will require training of the staff to an adequate standard to operate the User Interfaces (UIs). Virtual environments used for training will need to meet a standard that provide both an adequate training environment, and also an adequate testing environment. The EVW prototype 3 trials have been prepared to do this.

Surgical simulation is still in its infancy. So too is its relationship to the evolving Virtual Orthopaedic European University (VOEU) Project. Simulation and other such surgical systems will benefit training through Inter-Active Learning, and certain sub-skills of surgery explored in the SOTON Simulation are incorporated into a broader educational package that realistically maps the roles of the training surgeon in accordance with their ‘educational ontology’. This is to determine the capabilities of the Surgical Skills Simulation, through psychomotor performance vs. skills Levels testing.

The tasks of the EVW prototype 3 trials, based upon the EVW prototype 1 simulation trial results, support a training deficiency diagnostic approach and indication of the information-processing and psychomotor skills systems failures. Ultimately the simulations may then be developed, based upon accurate quantitative information regarding surgical performance metrics. This depends upon the tasks demanded of the University of Southampton simulator by the trainees in the EVW prototype 3 trials.

Such core tutorial skills to be trained focus upon identification of structures whilst navigating through a three-dimensional, '7 degree-of-freedom' space, and analysis at this stage of development is based upon user questionnaires that are aimed at the evaluation of the broad spectrum of surgical simulators. This is a subjective analysis of the system by trainees but still represents a basic evidence base upon which to build.

The comments recorded by the user group in the 1^{st} EVW prototype simulation are part of a process of evaluation evolution. The 1^{st} EVW prototype simulation trial results led to further design team group discussions. The results suggested that such browser-based interfaces are to be built into educational environments to assist the control of training simulators, and so this approach has henceforth been adopted. The EVW prototype 2 trial of the 2^{nd} generation systems did not proceed and so plans were established to create a further prototype incorporating the principles learned in *the 1st and 2^{nd} prototype simulation* for the final trial of this thesis.

The UI Design Strategy evolved also. By way of reference to standard models for User Interface (UI) design, this section included the analysis of user-centred design and its significance to providing interfaces dedicated to surgical users. This is essential for developing interfaces for surgeons. A key part of the usability trials of future surgical educational systems is the assessment of the interface and its integration into such systems.

A core design principle is the separation of structure and content. This is in part to ensure that the demands made upon the user are achievable. Once established, the interfaces will dictate staff training protocols as users will be obliged to adopt such systems. The content should be able to be transferred from system to system – by virtue of being contained in 'learning objects' categorised by specific metadata.

Usually, the system internals are designed first (entities, events, functions, etc) and the detailed user interface added later. Modelling users and their tasks, environment, and usability requirements earlier in the development makes it possible to develop an abstract definition of how the user thinks of the system (Users' Conceptual Model). For the Surgical Simulator Shoulder Arthroscopy Tutorial the software application allows the user to adopt a user-centred perspective upon the environment. This will facilitate integration of learning experience with previous clinical experience.

The research therefore led to the development of tools for Visual Integration. This supports the basic user profile. Ultimately the development of the spectrum of devices that can process and deliver data to users via a shared User Interface will convey a uniform application environment in the domain of healthcare information assimilation, analysis, and delivery. An XML proto-standard originally referred to as Orthopaedic Surgical Mark-up Language (OSML) was structured to support this principle. This has now for the most part been succeeded by the development of XML schemas to support the data structure of specific functions of the system.

Users are banded into groups of stakeholders, including basic surgical trainees (BST) and higher surgical trainees (HST). These are the target groups for the trials of the demonstration prototypes. The system also caters for providers such as editors who contribute material *e.g.* surgical knowledge engineers, surgical curriculum developers, and surgical researchers. This ensures that the experiments can control the stratification of user experience and thus the validity of comparison of different user groups. With such an underlying architecture developed, the next chapter goes on to describe the discrete tools within the shared working environment.

6 *The requirements and design of a virtual university infrastructure for learning orthopaedics*

6.1 Overview

The initial simulation used the concept of video-based simulation in the context of a discrete task-training scenario. In fact the world of surgical practice involves a multitude of such tasks that need to be managed often simultaneously. Success has been had on the undergraduate level with the introduction of *'mock-up'* operating theatres for the training of emergency procedures. These training environments offer a basis for developing teamwork training but are extremely expensive and so are only available in a few very specialist centres and are thus limited resources with respect to time and geography.

These skills labs are of value for *'on/off'* skills training and for competence assessment but cannot reasonably meet the need for *'just-in-time'* training, where the user needs to gain skills or access to specific information at short notice. In particular, a flexible distance-learning tool that has ubiquitous access and offers users material at an appropriate level of experience and relevant to their specific educational and clinical commitment is needed. This led to the VOEU project. The aims of this are to ensure access to all relevant educational material with the development of tailored courses for individuals, offering a customised view of the multimedia knowledge to which they refer. This links into issues discussed in the previous chapters by ensuring that the surgeons are able to incorporate the learning material into their working practice. They can do this by accessing material only relevant to their present situation, through the use of structured learning agreements. These are to be used as the framework for building a surgical ontology.

6.2 Introduction

Surgical Training and the tutorials to support it are about 'knowing what to do and being able to do it'. This means appropriate *a priori* knowledge, analysis of the situation and the ability to perform tasks safely and efficiently, with awareness of the appropriate feedback mechanisms. By reference to standard models for User Interface (UI) design, including frames and style sheets, this section includes analysis of user-centred design and its significance to providing interfaces dedicated to surgical users. This links to the principles of a Multimedia Educational structure, where introducing new media can bring with it greater complexity, and thus the risk of detraction from the content. Comparison with other multimedia visualisation techniques, such as video, 3D modelling and interactive video, is made and the significance of the bandwidth and delivery medium is discussed.

The interfaces influence *Staff Training Protocols*, so that setting up new systems will require training of the staff to an adequate standard to operate the User Interfaces. This requires the design and evaluation of an appropriately constructed UI that is intuitive. This is discussed in relation to the 1st prototype simulation Shoulder Arthroscopy Tutorial and its associated protocol for evaluation.

The progressive honing of separate sets of skills consolidates surgical skills. This creates skills 'islands' of knowledge. The structure and content of the shoulder arthroscopy tutorial is explained and expanded upon.

This section considers both the training material and how this is integrated into an interactive multimedia educational orthopaedic module software application to support the aim of embedding the simulation into the learning environment in a meaningful way. In the 1st prototype simulation this was achieved by providing a web-compatible interface written in Java supporting upgradeable platform-independent UI. This was intended to minimise the problem of obsolescence of applications. With many new formats evolving more rapidly than the standards that may unify them, the decision was made to adopt customised off the shelf (COTS) components with Microsoft Internet Explorer (v5.5 or later) providing XML support. This is market driven, and risks developing platform dependence. The issue, and its significance, is explored below.

6.3 Requirements of the virtual university infrastructure

To lead the way for setting suitable standards for the trainees and trainers, the aim is to establish a code of practice for the use of new surgical technologies, including simulation. This will be achieved by analysing how the instruction can be implemented (the delivery mechanism) in parallel with qualifying how to train the trainees and trainers.

This will outline the specifications for equipment that they shall use. This will maximise the opportunity to look at the development of the field, and ultimately to establishing criteria for the CE mark. Such a process of validation and verification (V&V) needs to address the domains in Table 6.1:

	Component to be tested	**Evaluation tool**
1	Engineering and Networking	Usability (Questionnaire A)
2	Surgical Systems	Virtual Environment (Questionnaire C, parts A, I, P)
3	Training by Inter-Active Learning	Education (Questionnaire C, parts E, S)

Table 6.1 Master test plan for the EVW prototype 3 simulation

The key areas of technological development in simulation and engineering are evolving ‘hand in glove’. Based upon the views of the author, Dr P.Brett and Dr E. Holler (120;121), key areas that require scrutiny by the proposed Verification and Validation processes are defined.

Engineering planning should take advantage of developing broad bandwidth Information Technology infrastructures to provide a uniform quality of service for inter-active learning. What different types of man-machine interface there are should be outlined, and what they portray to the user.

Considering how simulation performance might be assessed, there is the potential to evolve a database of average performance that can be used to compare individual surgeons to a standard, and this can then validate the simulators themselves. This poses the question of what new solutions for the development of surgical presence (sense of immersion in a VE) are on the horizon and how these might be integrated.

Educational consequences of inadequacies in the simulated environment are important. This relates to the likelihood and consequence of errors in the real environment, leading to entry-level standards for individual sub-skills training in simulation. These new criteria may be integrated into a coherent set of guidelines for the introduction of these new technologies into the operating and training environments within the existing European Health Care infrastructures. This will require clear definition of what simulators should be used for, *i.e.* which surgical sub-skills can be assessed. Ultimately this leads to implementation planning, who will benefit, and how, including the manufacturers of the sophisticated tools.

The author proposes that such guidelines be established for systems that are going into clinical training environments. This is discussed in more detail in Chapter 9.

6.4 How adaptive hypermedia will change the role of the educational portal environment

Rather than simply accessing the educational environment (portal) as a user, the portal becomes the extension of the user through expression of their preferences. The role of the individual dictates the perspective upon the world. This will incorporate agents for the management of new tools and new roles that the individuals will adopt. The personal portal will become an essential appliance in the same way as the shoes we wear allow us to interact with the physical environment. It would be possible to walk everywhere barefoot, but probably not as comfortable! In the same way we need to ensure that clinical users accept the free flow of information around the world in which they work, and allow others to access their information resources and ontologies freely, once permission is granted.

6.5 Structural ontology of VOEU

Key to the use of such virtual university working environments is the flow of information. This relates to the role of VOEU and how it can accommodate the ethical intentions and requirements of the clinical educational process. The flow diagram in Figure 6.3 outlines the VOEU processes as they relate to the underlying educational process. The management of simulation integration must integrate with this infrastructure. The infrastructure is modified after the efforts of Tim David *et al* (122).

6.6 Pedagogy

The pedagogical strategy is based upon the creation of this novel computer *case-based learning* environment, which aims to maximise the relationship between the pedagogical approaches adopted and the tools and resources available to support them (Figure 5.1). This builds on current thinking in educational research on pedagogy in terms of learning being situated and authentic, with learners adopting an active and constructive approach. In particular it builds on the problem-based learning literature, constructivism (123;124), communities of practice (125), situated learning (126-128) and activity theory (129). The pedagogical strategy thus aims to create an environment that allows the different benefits of these pedagogical approaches to be made explicit. Guidance with exemplars must be given of how problem-based learning can be used in conjunction with collaborative learning. Through the use of the case studies, the *DRJ* and the communication environment are therefore included as part of this learning environment. These learning components are discussed in detail in Chapter 7.

The framework is designed to take the course convenor through the thought processes of a course. Different media types may be assessed and the different educational interactions supported. Resource issues and local constraints will dictate the learning agreement boundaries. This toolkit for individual learning agreements maps a new individualised course.

The toolkit models instructional design (130). With tools provided at key stages along the path, a customised course is developed, leading to an individual profile for the trainee surgeon. These individual tools are designed to help the user access a knowledge base in order to make informed decisions. The format of toolkits allows users to access components in a standard, linear fashion, or to explore in detail if the users' expertise is stronger in some areas of knowledge than others.

The pedagogical strategy is thus eclectic. It is designed to invoke active participation making use of the multiple resources available in the educational environment. In addition, it is designed so that the users who are motivated to learn about a topic can do so independently, by searching for, evaluating, and using authentic information. This proposed learning experience endeavours to mimic real life, targeting the learner as the information seeker and interpreter who constructs knowledge through problem solving with information tools.

The advantages of this trainee-centred approach to learning include the promotion of the development of thinking skills such as problem-solving, reasoning, and critical evaluation. Such an infrastructure should improve the research skills of the trainees, supporting the research work of the teams. It also offers flexibility accommodating different learning styles. The academic (declarative) work the trainees carry out is correlated with their procedural work upon key skills (*e.g.* the surgical logbook).

6.7 An integrated approach tool building

Practical software engineering requires experimentation. Figure 6.1 suggests exploring the area where pedagogy, tools and resources interact.

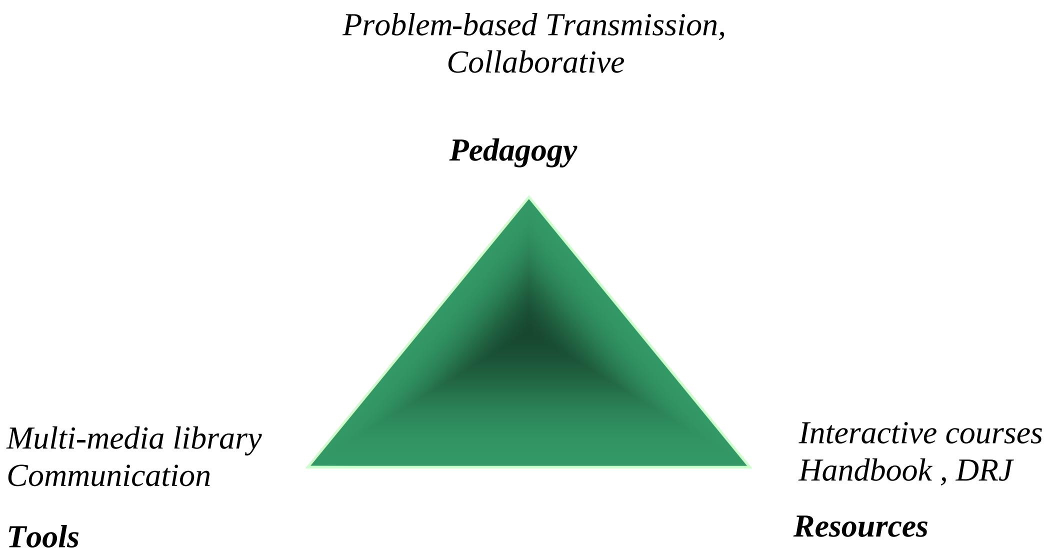


Figure 6.1 Relationship between pedagogy, tools, and resources

This approach generates lists for the syllabus, curriculum and content. The granularity of checklists is bound to vary from topic to topic, but by using the established syllabus it is possible to bring the relevant resources to fit into the overall pedagogy. This allows it to be built into the framework and accommodate different curricula, syllabi and learning tools, such as case-based learning in the examples demonstrated. The stages of development are outlined in Figure 6.2.

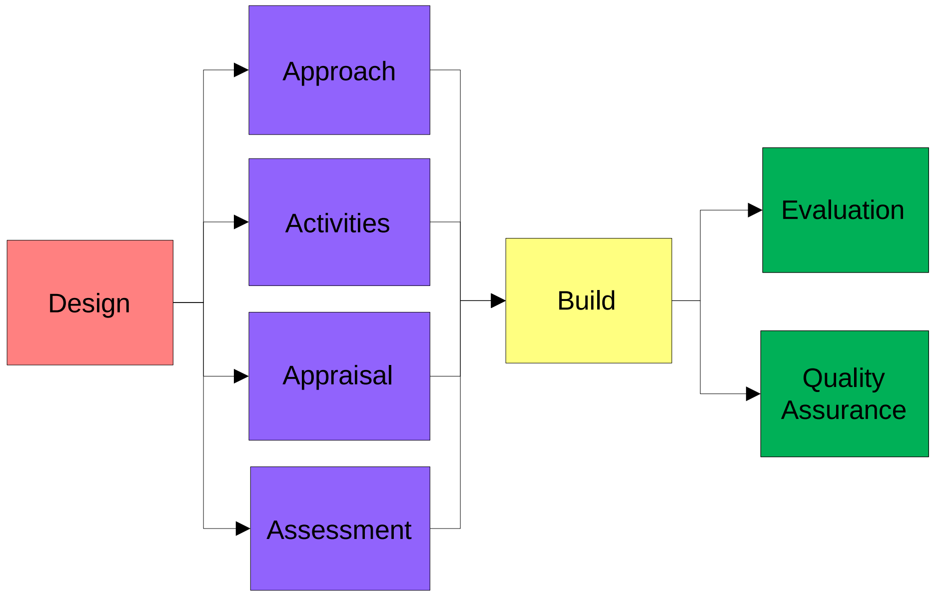


Figure 6.2 Educational Tool Development Process

6.8 The on-Line VOEU syllabus for orthopaedics

The VOEU project included many specific courses beyond those used to demonstrate integrated simulation, which are based upon the syllabus. The basis of the educational contract was to bring the educational syllabus and the individual learning agreements together. This provides a two-way proposal for the tutor and trainee to agree initially and then review as the clinical post progresses. The aim is to see the areas of overlap between the trainee's learning needs, the tutor's ability to focus upon specific clinical areas, the resources of the clinical department and case mix to provide relevant learning material and clinical experience. This will map on to the personal profile of the individual which links to the ontology (Figure 6.3), selecting relevant points of the syllabus, integrating the contract and continuing completion of tasks.

The syllabus reflects one aspect of the modern surgeon's role, that of being a conduit for knowledge, acting upon it or passing it on to others so that they may execute actions. The learning agreements reflect a formalisation of the relationship between tutor and trainee, and collectively between training organisation and trainees in general. Teaching takes time and consumes considerable resources, so its justification is being evaluated constantly with the audit of its delivery being managed by learning agreements.

Learning agreements broadly coincide with the purposes of a structured education system. There are ways in which this is different, particularly with regard to meeting established standards. The benchmark for the development of the syllabus initially and the learning agreements is the medium-term need for individuals to pass the certificate of completion of specialist training (CCST) exam (131).

This approach will enable:

1. Preparation of an outline course infrastructure for trainers, such as course convenors or instructors and the tutors of individual trainees for specific courses or periods of in-house training (posts of usually 6 month's duration).
2. Curriculum development that is dynamic, building in units that can be combined and decomposed in meaningful ways. This will be mapped onto the syllabus and structured to suit the individual learner's needs. This is currently under review by the curriculum development committee (132) at the Royal College of Surgeons of England.

3. Documenting and recognizing the completion of existing or new learning and performance objectives developed as part of VOEU, such as in the VOU3 trial.
4. Education, training, and learning organizations, to monitor an individual's progress as related to covering the syllabus (39).
5. The necessary security and authentication (including non-repudiation) for the distribution and use of Learning Agreements in particular – in accordance with the Data Protection Act and the laws of consent regarding patient information.

Figure 6.3 The VEOU educational workflow

6.8.1 Surgical ontology (133) and the learning agreement

The surgical educational ontology is the basis for the learning agreements, but there is in addition to this the need to accommodate the organisational approaches of different individuals. By formalising the process, the *Learning Agreement ontology* can then be standardised and provided as a part of the pedagogy for the development of courses so that these may be integrated with the trainee's specific needs and the trainer's (clinical tutor's) ability to accommodate the specialised learning needs of the individual within the context of the most suitable caseload and experience within the setting of the clinical post. The process involves negotiation (Figure 6.3) and permits non-repudiable record keeping as suggested above. This allows tutors and tutees to relate progress to specific events in clinical logbooks, as seen in Figure 7.9.

6.8.2 VOEU architecture and implementation

Digital technologies support a working environment for familiarisation with new surgical procedures and the management of clinical case audit. To enhance access and control of these virtual workspaces the eXtensible Mark-up Language (XML)-based interface links the educational environment of the Web-Based Training (WBT) scheme to the clinical data collection from ongoing trials in Orthopaedics, as mentioned in Chapter 9 regarding the DRJ.

The VOEU infrastructure encompasses the normal working environment of engineers and clinicians using computer-assisted surgery tools adopting dedicated interactive media. An XML-based interface enables users to communicate using material mediated for their specific needs allowing adaptive media, based upon user experience and knowledge base. This combines the declarative (factual) content with feedback from the clinical (procedural) case-based training and evaluation environment. By using the XML interface, one can cater for the different user hardware and software resources, media content, and language of presentation, incorporating virtual environment simulations.

6.8.3 Adaptive hypermedia architecture

When constructing the framework for the VOEU application, an information model was developed to feature three types of 'separation', namely: the separation of presentation from information (content), the separation of links from content, and the separation of logic from information. These three features were taken into consideration in the design in order to reduce the authoring effort by re-using the resources and to facilitate the maintenance of the application.

Techniques employed were link databases or *linkbases* (134) to store links separately from the documents to which they refer. This separates links from content. Links are embedded into the later simulation models to ensure bi-directional searching between video and adaptive hypermedia content.

The idea of generating *associative links* comes from the usual phenomenon of users browsing in the current context to find more information about a particular concept. For example, when users are reading a document or solving a problem that concerns concept 'X', links to other documents are suggested, which they could follow. Hence, these links are called associative links because they associate two or more documents based on a common concept.

Associative links are stored in a linkbase where each of them is indexed or retrieved based on a keyword or concept. From a learning perspective, this link type is used for goal-directed learning where the user will search on a concept to find a series of related documents. It is our intention to develop this and also the user-prioritisation links based upon the priority of material in the actual implemented version being produced currently at http://vou.soton.ac.uk.

6.9 Prototype four

This involves the methodology development for each version of the simulator, building upon previous experience. It has itself evolved so that for the 4th prototype, the virtual university will incorporate the use of threaded discussion groups for the process of building any future designs in a safe, secure, accessible environment. This is a first step to building naturalness into a conversational multi-modal interface (135) and an integral part of the planned development of collective working environments for scientific research in the broader domain extending beyond orthopaedics. It allows for the linking of data collection and results analysis using the virtual observatory and dynamic review components of VOEU.

6.9.1 Issues involved in evaluating virtual environments

The thorny problem of re-certification may be made more palatable if surgeons are asked to spend an enjoyable day in the surgical simulator every six months, as airline pilots do, for re-certification, rather than be forced to undertake written examinations every five years. This process of revalidation is still being addressed by the policing organisation for UK doctors – the General Medical Council (GMC)[26] (www.gmc-uk.org). Standards should be set, with increasing CME credit being awarded for ascending complexity of simulators, as occurs in the airline industry. Until a process is agreed, the prototype of simulation need only support this principle through links to the Virtual University infrastructure.

6.9.2 Relationship to the evolving Virtual Orthopaedic European University (VOEU) project

In order to provide a coherent strategy for the development and evaluation of the surgical training infrastructure a concerted effort is required to develop and uphold an ethical code. Over-regulation may strangle the potential new technologies, restricting both their use and their development. This is a fine line to walk, but should be achieved through the preparation of minimum standards for the accuracy and performance of tools. The VOEU Project achieved this. The cross-disciplinary project received European funding under the Information Society Technologies directive aimed at enhancing the training of orthopaedic surgeons using digital technologies.

6.10 Basic surgical computer literacy

To introduce a VE-based surgical simulation requires a population of computer-aware individuals who are willing to adopt such a training system. Better to have co-operation than coercion. In order to ascertain where the population is on the computing 'learning curve' and to make the most of developing intuitive interfaces relatively familiar to most surgeons. An audit was conducted of surgical computer performance and training. The questions included:

- How much training had surgeons received?
- How much investment were they prepared to make *e.g.* purchasing their own equipment?
- What would be the best approach to getting better computer training, and hence literacy in the surgical population?
- What would be the most useful skills to teach surgeons?

[26] More will become apparent as the medical profession accepts the need to build this into their continuing professional development plans. The aim is for the VOEU user profile to support this. More information is available about revalidation for doctors on-line at www.revalidationuk.info/.

- Would they have any experience with which to compare a simulator?
- Will future training programmes using remote access to teaching material be deemed acceptable by trainees with as yet no such infrastructure being in place?
- What are the results of past and current VE evaluation efforts?

The survey was limited in time and focused for the sake of the pilot study for the role of the shoulder arthroscopy simulator and its interface, and so this limited the questionnaire to questions in Table 6.2:

Question	**Response Range**
Have you used a Surgical Simulator before?	Yes No (if Yes, where? and which one?)
Would Teleconferencing rather than travelling to formal teaching be acceptable? And if yes, for what percentage of the sessions?	Yes No 10% 20% 30% 40% 50% 60% 70%

Table 6.2 Previous virtual working environment experience

6.11 Conclusions drawn from the 1st prototype simulation trials

There is much headway to be made with the training of ICT skills in the surgical population. Whilst the author is aware of senior clinicians who are still computer illiterate at the time of writing, the majority of trainees are now using the applications for supportive clinical work even if not everyday care delivery.

6.11.1 Significance of the 1st prototype simulation results

This state of affairs is quite striking, in a population that the UK government intends to have 100% on-line by 2010, with the planned introduction of revalidation likely to require significant audit data collection, especially when one considers that the population studied have all had a university education. The aim of the *ICT Skills for Healthcare Professionals* course is to bridge the gap between the *ICT Haves* and *ICT Have Nots*. This is proposed as an addition to the ‘on-line’ instructions to provide a ‘help’ service for VOEU and to structure a definitive course for those who wish to use it. This will be vital to the uptake and acceptance of the Virtual Orthopaedic European University (VOEU) within the orthopaedic community. On a positive note, the willingness of the population to evaluate and embrace new technologies is encouraging. The majority wishing to use teleconferencing facilities for nearly half of formal teaching sessions (46% based upon the average preference for the group) highlights this.

6.11.2 Potential sources of bias

The most apparent biases are temporal and geographical.

- **Temporal:** The study was carried out in 1999 and represents the situation before the start of the VOEU project. It was necessary to collect the data again from trial candidates as they sign up to VOEU participating in the beta-testing phase as part of *the 3rd prototype simulation.* This is to get a temporal perspective in the uptake of digital technologies, and then to use this to target areas of limited expertise through future modifications to the educational courses.
- **Geographical:** This population were all from one region of one country. As part of a European trial, which will generate the main study of the thesis, the aim is to address this issue by developing the questionnaire as part of the registration process of VOEU, so that it will be possible to clarify which areas of Europe are more able to adopt VOEU, so that the policy for the Dissemination and Use Plan (DUP) can accommodate this.

6.11.3 Conclusions to be taken forward to the VOU3 simulation study

The following matters need to be addressed:

- Provision of a broad based ICT training resource for surgeons through the delivery of Multimedia Educational Orthopaedic Modules (MEOMs) for the 'Basic Computing For Surgeons' course.
- Provision of an ICT 'help' service for surgeons through the delivery of Multimedia Educational Orthopaedic Modules (MEOMs) for the Basic Computing for Surgeons course material being accessible through the search engines in VOEU.
- The need for updating this data resource on a regular (possibly biannual or annual) basis, which can be achieved through the DRJ component of the virtual observatory. This system will allow for ongoing results reporting, which is being developed as a part of VOEU.
- Integration of Flexible and Distance Learning (FDL) into surgical training using teleconferencing, which is beyond the remit of this thesis, but will be discussed in VOEU.

6.11.4 Questioning the validity of the 1st EVW prototype simulation results

- **Are the results internally valid?** The results have been checked and are arithmetically accurate. They represent the views and performance of around 90% of the population who are eligible members of the user group.
- **Confounding (Distortion of apparent effect):** This can be measured and controlled for. In light of the above figures, there is not enough data to stratify to eliminate all potential confounding variables, and so stratification by subgroups has to be regarded with this *caveat* in mind.
- **External validity – Generalisability:** Are the results valid for another population *e.g.* your local population? With the United Kingdom, due to safeguards in the maintenance of training standards, the results are likely to be generalisable but may still indicate a national geographical bias.
- **External validity – Importance:** The nature of the work analysed implies that external validity is vital to the acceptance and uptake of such systems.
- **External validity – Significance:** This matter has already been addressed in Chapter 1 reflecting the history of surgical simulation and its motives, where the current social, political and medical motivation has been explored.
- **External validity – Importance vs. Significance:** The work is both timely and necessary. Continuing within the framework of a major European (5^{th} framework) study adds external peer review, particularly concentrating upon the educational aspects of the work and credence to the evolving methodology for modelling and evaluation in surgery.

6.11.5 The 1st prototype simulation (EVW pilot study) evaluation methodology

The results were evaluated using the NASA-TLS (136) scale for the assessment of tasks, which allows users to rate tasks on a linear scale from 1 – 7 for complexity or difficulty. The mean values for the focus group are then expressed as a percentage. This method was extended to the main study, alongside the constructivist approach expanded above, in order to provide a comparison of the evaluation methodologies.

Redesigning the AIP, and other questions: Much of the original work of evaluating VEs was described in Barfield's book (137), and the general questions (many used by NASA) have been modified to assess the specific environment's needs. This forms the foundation of the reductionist methodology since it extends the Zeltzer (AIP) cube. As detailed in Chapter 4, the AIPES modification should more readily reflect the aspects of integration of the simulation that are not covered in AIP.

6.12 Which surgical sub-skills are being assessed?

From the above results it is apparent that sub-skills explored in the 1st EVW prototype simulation trials demonstrate that this simulator is in its infancy, monitoring only certain key sub-skills, but the data described below forms the basis for an evolving platform for the measurement of surgical performance.

Currently there is limited haptic feedback of Minimal Access Surgery (MAS), but this will change (MAS poses the fewest technical challenges of any type of surgery to model in computer simulation). This was not formally evaluated as part of the 1st EVW prototype simulation pilot study. Those sub skills marked in bold in Table 6.3 below have been investigated using the EVW system, focusing upon visual accuracy of the VE.

Other authors (67) have proposed that psychomotor performance standards are constructed by testing a population of surgical trainees and consultants using the following indicators:

- Task Analysis - Input of Subject Matter Experts (University Teaching Team)
- Determine the capabilities of the Surgical Skills Simulation
- Simulator-Based Trials with novice and experienced Surgeons
- Psychomotor performance vs. Skills Levels testing.
- Skills Clusters ~ criticality rating based upon the simulators' testing resources.

These can be changed as time goes on and results suggest new sub-skills can be tested in line with parts of the system in accordance with the structured surgical course model. Performance standards guidelines (the tasks to be tested by the simulator) may then be outlined for the main study.

Phase of procedure	Sub-skill	Training Level	Media
Pre-operative	Preparation Consent	A B	Text Text
(Peri-operative)	Marking Local Anaesthesia Sterility precautions	A A A	Text Text Text
Operative	**Orientation** **Navigation** **Pattern Recognition** Recognition of Pathology	**B** **B** **B** C	**The 1st prototype simulation +2** **The 1st prototype simulation +2** **the 3rd prototype simulation** Version 4
(Peri-operative)	Record Keeping	A	Text
Postoperative	Explanation of Findings Wound care Fitness for discharge Out-patient review	A A A B	Text Text Text Text

Table 6.3 Sub skill associations with stage of procedure

6.12.1 VOU3 simulation study tasks based upon version 1 results

Ideally the aim is the diagnosis of Training Deficiency, prior to a *'just-too-late'* deficit in experience when dealing with clinical situations. The system should be able to identify deficiencies in the pre-operative and operative skills. It should be able to provide the users with post-operative feedback that is objective, measuring targeted skills.

This corresponds to Surgical Behaviour, but in order to measure it, a diagnostic component needs to be included to track the performance error to the **root** cause. The process to track error in decision-making is complex and may evolve using the Dynamic Review Journal development as part of VOEU, though time and staff resources are likely to relegate this task to 'future work'. If resources permit, then the paths of critical decision making to be tracked are in information processing, cognitive ability and psychomotor skills.

6.12.2 Discussion of future needs for EVW simulation in light of assessment.

The aim is to enhance the iterative design process of the shoulder simulator by providing a formal ‘Future Needs Assessment’ for ‘Minimal Access Surgery’ outcome document, iteratively, to help provide a framework for the next generation design specification. The simulator may then be developed based upon accurate quantitative information regarding surgical performance metrics. This will be based upon the *ePrints* publication process generating dynamic results of the usability trials that can then be updated. 1st generation (EVW) would need to capture time on task, and test sub-skills: some skills are not simulated. Proposals for parameters are outlined in Table 6.4:

Sub-skills identified by:	**Task Analysis**	**Expertise**
Skills Simulated	Skill Clusters	Criticality Difficulty Necessity to task Performance
Quality of Simulation	Sub Skill areas:	Spatial Relations Planar Motion Rotation Retraction Planar Motion* Two-handed spatial relations*

Table 6.4 Proposed sub-skills vs. task analysis (not tested using this simulator)*

This is however an unrealistic set of parameters to easily collect data on easily. An evaluation (validation) methodology needs to recognise the perceived value to the end user also, and at the time of the 1st prototype simulation trial this was the only set of metrics that were easy to analyse. Other spatial co-ordinate data have been collected but will not be analysed for this thesis.

6.12.3 How to achieve the goals of reliable metrics for evaluation as part of the 1st prototype simulation trial (EVW)

Psychomotor performance can be captured as raw data, providing the ability to record accurate measurements of performance. The difficulty is converting this into meaningful information for comparison of trainees to reliable benchmarks. Whereas descriptive statistics, such as comparing the confidence of a surgeon to perform a task with their grade and experience is perhaps of some significance since it displays trends, the important values of quantitative metrics are not addressed (see Figure 6.4).

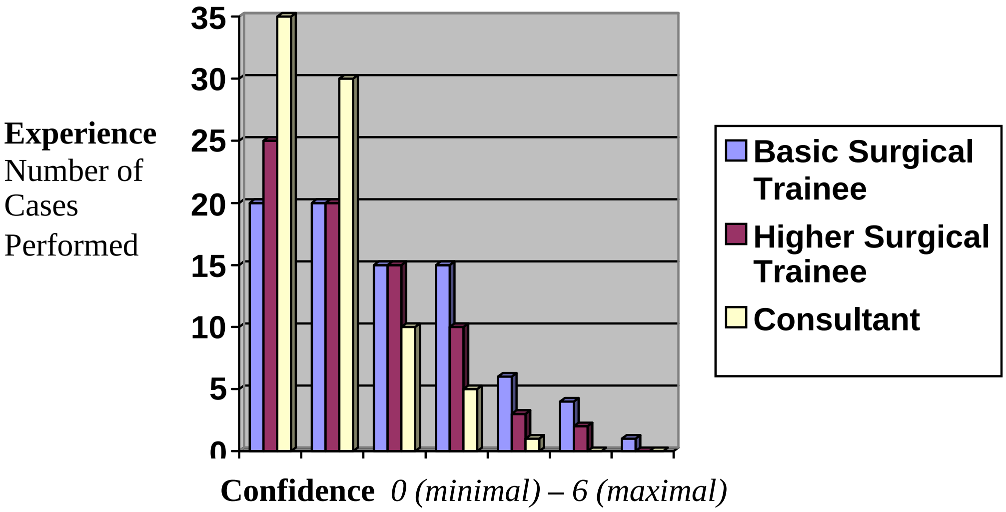


Figure 6.4 Experience vs. confidence in surgeons using the 1st EVW prototype simulation pilot study

The realism of the simulation and the context of the training scenario should be assessed during the AIPES evaluation of the simulators in *Questionnaire C*. An example of this is when video databases are used. Fundamental to the concept of presence *i.e.* to ensure adequate images are recorded in the video environment to minimise the effect of the granularity. This is one of the factors to be considered regarding the level of detail in the simulation. To understand the restriction of the numbers of frames necessary when using a video-generated database, the author would refer to other work in the public domain (138) for more information. Whilst the work so far demonstrates ways of assessing these particular metrics for simulator performance and accuracy, they do not reflect the need for measures of educational value. This is in effect a user-led evaluation process rather than a technical one, building up a catalogue of user-dependent tasks for the simulator.

6.12.4 Tasks demanded by the trainees in the 1st phase simulation pilot study

Realistically, the analysis of such level of detail is beyond the casual user evaluation and would require formal (funded) task evaluation. A simple hierarchy of tasks has been proposed before the pilot study outlined in Table 6.5. This may be scored subjectively using a qualitative nominal scale, where the complexity rating of * represents a simple process and a rating of ***** represents a more complex process due to increased number of required frames:

	Training Procedure	**Video Worlds**
1	Steering in an 'unfamiliar' environment (orientation)	*
2.	Navigation	*
3.	Pattern Recognition	**
4	Triangulation	**
5	Pathology Simulation	***
6.	Procedural training	*****

Table 6.5 Training procedure vs. task complexity

This suggested that the 1st generation pilot study version of EVW (tested in 1997) was able to address the basic tasks 1 – 4 but was inadequate for tasks 5 and 6.

- **Identification of structures:** This is calculated as a percentage of the peer group for pattern recognition performance for certain tasks such as identification of the long head of biceps, or the anterior Gleno-humeral ligaments. It is simple to measure since the track of the user will provide a spatial coordinate that can be predefined as a series of locations where the appropriate information can be harvested.
- **Time to complete tasks:** This concept is described in more detail elsewhere (36) but so long as tasks are not deemed time critical, then this is less important.
- **Subjective analysis of the system by trainees:** This is not to be underestimated as a source of information for feedback into the design process of shoulder simulation. This is considered a 'class C' evidence base. The comments recorded by the user group in the pilot study related to the Java interface were that it was *'Essential'*, with *'A lot of*

investment necessary and needed' One user commented that *'Java appears slow at present'*.

These issues were addressed in the next iteration of design. Forms for comments will also be included within the feedback process, allowing unbounded text boxes for comments. As will be seen below, Java elements for control were replaced in the 2^{nd} EVW prototype (University of Bristol) system by C++ components and then for portability in the 3^{rd} EVW prototype (University of Southampton) system by using intrinsic COTS controls (*QuickTime™ 6.0*).

6.13 Implications for the wider issue of simulation in surgery

The vocational training of anyone whose work involves clinically invasive procedures requires the best from that individual with regards to decision-making and psychomotor skills Thus the development and monitoring of these is vital. Ultimately simulator-based skills training will become integrated within the framework of clinical governance (the *quality assurance* of clinical service delivery). The methodology of the *1st prototype simulation* task analysis for this trial was based upon the practical limits of what can be achieved with current available equipment, and also the expert analysis of tasks (61;139). The results of the *1st prototype simulation* trials were inconclusive with a small number of trainees, as would be anticipated in a hypothesis-generating study with a small population. A larger number of trial candidates are required. This led to the Virtual University supported approach in the VOU3 trials.

The tutorial system needed refinement as well as the simulation system, and so parallel development of the different domains (library, virtual classroom, discussion groups and the dynamic review process) meant that by VOU3 the user population and the tutorials were supported by a fundamentally different infrastructure from the original Exeter Virtual University of Orthopaedics1 (EVOU) shoulder arthroscopy tutorial system developed for *The 1st prototype simulation.*

6.13.1 Process of evaluation evolution

The pilot study results led to further *Design Team Group Discussions.* These were similar to Focus Group Discussions, but included involvement of the interdisciplinary design team of engineers and medical personnel. Design Team Group Discussions lead to further design modifications to both the methodologies for the software and the *User-based evaluation.* This involved plans for the controlled testing of the system while it is being used for specific tasks of the procedure. The *Questionnaires* for collecting the user's experience supplemented by interviews and other discussion methods were thus improved and updated as a consequence of the evaluation. Such a process for evolution of the *Questionnaire A* methodology (Chapter 8) usability group runs in parallel with the iterative design process for simulation products.

6.13.2 Review of the use of Java as an interface.

The primary purpose of the Java code was to provide a channel of communication between the two platforms of the tutorial and the simulation achieving integration (in the form of embedding), providing different parts of the system to the end-user. Use of this should not require skills beyond the basic information technology skills of an average surgeon. The level of this was confined to the user profile analysis in *the 1st prototype simulation.* The assertion that the interface can be built into the conventional User Interface (UI) is aimed to assist the uptake of the technology. The development of the user interface is likely to progress in two directions:

1. ***Improved communication between machines*** using the developing agent technologies (often referred to as software robots (autonomous software applications), which is yet to be implemented.
2. ***Increased complexity of the interface*** through improved functionality. This has been implemented in the VOU3 using XML It should, where possible, abide by standards being evolved for this domain by the international organisation Health Level Seven (www.HL7.org). This will ensure greater compatibility between the interfaces that are used for the control of surgical procedural trainers (simulators).

There were potential pitfalls using Java that precipitated the change of approach with respect to tutorial design and integration. The most significant of these are:

Security: Although the system is stable and likely to display good survivability characteristics (140), the open nature of the language and its potential application for providing communication portals between the machines can provide a vehicle for corrupting software (viruses) also. This application was designed only for operation between trusted machines. *Remote Method Invocation* (RMI) may help to overcome this problem using Java. Custom socket factories (software tools that assign the connecting sockets between computers) can allow a remote object to specify the protocol that RMI will use for remote calls to that object. RMI over a secure transport (such as SSL) can be supported using custom socket factories. However the bigger picture calls for robust secure extranet services and so a platform was developed to meet this demand.

This in fact runs JavaScript over http using the *href="javascript:open_movie('shflick1_u.mov',360,288,280500)"* command inside the secure extranet area – a command invoked from;

http://voeu.ecs.soton.ac.uk/voeu/library/doc.aspx?docid=13 .

Speed of processing: the *1st prototype simulation* pilot study user group highlighted this in subjective comments. There are certain applications for which the Java interface is not the most appropriate tool. For the development of complex simulations, *C* and *C++* provide a fast and effective language for coding. Designers need to keep this fact in mind when considering multi-user environments where critical action sequences must be maintained in the simulated environment. Java has a role for the control process at present but not for the simulation engine itself. This was indicated by the pilot study assessment. By changing the VOU3 model that limits interactivity it is possible to greatly improve access, and by adopting video streaming technology it is possible to ensure that this runs upon lower specification platforms, as is likely to be found in most hospitals.

6.14 User interface design strategy for surgeons

By reference to standard models for User Interface (UI) design, this section includes analysis of user-centred design and its significance to providing interfaces dedicated to surgical users. For everyone, intuitive interfaces need to be quick to read, assimilate, and interact with. A key part of the usability trials of future surgical educational systems is the assessment of the interface and its integration in such systems. A unified interface (the Visual Integrator – referred to by the Helmholz Institute[27], Aachen, Germany as VI) will assist course integration through *Re-utilisation of components* for easier course construction. This leads to the evolution of the structured surgical course model, where users are presented with a course that is viewable in the format of a standard recognised pattern of course design if users wish to adopt this approach. This therefore develops added value and 'Brand Loyalty' as the user becomes

[27] The Helmholz Institute, Aachen, is known for its expertise in ergonomics in the medical technology arena. For more information please review: http://www.hia.rwth-aachen.de

more familiar with the working system that presents material in an approach familiar to them.

6.14.1 Windows

The sub screens and specific toolbars of the future surgical educational systems should be defined. Toolbars are likely to include:

- **Journals –** In the form of the ePrints archive
- **References –** Reference Manager database system supports this.
- **Courses –** Structured in whichever format is deemed appropriate to the user and the situation.
- **Lectures –** Archive resources based upon the material already collected.
- **Workgroups –** Forums can be created inside future surgical educational systems. The stakeholder levels need to be adopted. This should be arranged, based upon a meritocracy. The moderator of the group can manage a list of suitable stakeholders. The current design of VOEU supports icons for the library, dynamic review journal, educational section, discussion forums, personal profile, and toolkit sections.

6.14.2 Icons

Stakeholder identifiers can be displayed along with icons for the branding links to key institutions. The same approach applies to journals regarding access and on-line registration available via the future surgical educational systems. This will require the ability to place appropriate icons into the browser windows with the permissions of the companies involved. Frequently used VOEU application icons are detailed in Figure 6.5.

Library – document repository

Dynamic Review Journal

Virtual Classroom (Question bank)

Discussion forum

Personal Profile

Toolkit – System administration tools

Exit (and Log Off)

Figure 6.5 Frequently used VOEU application icons

6.14.3 Menus

The menu hierarchy should be included within the design specification, but the middleware must be kept separate from the content. Publication channels for previously authorised material should be included. There should be no change in the authors' rights to use the material so long as dual publication within digital and paper media is prohibited. This is to be handled by the *ePrints* server, which adopts the principles of the Open Archive Initiative (OAI)[28].

6.14.4 Pointers

Currently no special pointers are planned covering the whole of future surgical educational systems. The adoption of the standard 'Windows' pointers is proposed. It is anticipated that in the next prototype the files will support colour coding for the representation of appropriate paths based upon whether documents are labelled as essential, important or further reading. This prioritisation will be based upon their user profile.

---

[28] Open Archive Initiative (OAI). This is an established standard for the interchange of information. See www.openarchives.org/. A list of registered sites is available from: www.openarchives.org/Register/ListFriends.pl

6.15 Multimedia educational structure related to the user interface

Making a good tutorial is a combination of writing a good book, making a good film and holding a good conversation all in one. The medium may be different but the educational principles are unchanged, as emphasised by a structured pedagogy. Introducing new media can bring with it greater complexity. The risk is of the media detracting from the content. For this reason the approaches to the User Interface (UI) design have changed as much throughout the project as the UI draft versions themselves. Originally driven by the quest for standards adaptation, somewhat rigid design methodologies were originally employed. The order was thus:

- Structured Systems Analysis and Design Methodology (SSADM)
- Graphical User Interface Design and Evaluation (GUIDE)
- Iterative Design (ID)
- Rapid Prototyping (Extreme Programming)

The reason for adopting this approach is that although overall the project work is relatively large scale, the individual components can be kept small. This facilitates a modular design forming the building blocks. Each may be designed, built and tested by rapid prototyping in order to alpha test the ideas and procedures.

This then provides feedback for the next stage of the design. The process to develop this approach is explained below. The evolution of this methodology more readily reflects the influence of several interrelated factors:

- Increasing expertise of the developers
  - Decreasing timeframe between prototype versions
  - Development of an established body of software.
- More relaxed attitude toward the design methodology
  - Broad range of options supported by the later versions
  - Separate structure and content.

6.15.1 Comparison with other multimedia techniques

Multimedia visualisation techniques, such as video, 3D modelling and interactive video, operate by activating a media file and provide primarily visual feedback. They differ in that they require various different drivers, and video offers limited navigability in 2D. Video also imposes restrictions upon interaction, and graphic generated images, unlike video, are not bound to have a fixed timeline.

The significance of the bandwidth and delivery medium is fundamental to the rollout plans. An exemplar is the low access to the system over a conventional 56 KBps modem. This means that 'entry' level access is going to be via so-called 'broadband' – 2 MBps bandwidth.

6.15.2 Interfaces dictate staff training protocols

Setting up new systems will require training of the staff to an adequate standard to operate the Graphical User Interfaces (GUIs), but this requires the design and evaluation of an appropriately constructed UI that is intuitive. The processes where human – application interface design is important, adopting the Graphical User Interface Design and Evaluation (GUIDE) approach, are described in detail in this section, though the emphasis is upon the user-centred design philosophy. This needs to be tailored and often reverts back to the principles of Structured Systems Analysis and Design Methodology, which are discussed first.

6.15.3 Tailored approaches to systems development using SSADM version 4

Where possible, the work of the EVW 1st and 2nd prototype generations used the Structured Systems Analysis and Design Methodology (63;141) to make it easier to adapt the system to other applications as it evolves. The VOEU stakeholder analysis adopted in prototype 3 provides a more detailed picture than SSADM of who the users are, what they are trying to achieve, and how they perform their tasks. This is considered essential for a user-centred approach to application development. Consequently the External Design included the Users' Conceptual Model, as indicated by the ontological issues raised earlier in this chapter. UI Design involved User Interface Prototype development and an Evaluation Report.

There was therefore a change from core SSADM in the way the External Design is created, evolved and validated. Usually, the system internals are designed first (entities, events, functions, *etc*.) and the detailed user interface added later. This represents a bottom up approach, so what the end-user will actually see is not known until later in the process. Unless Specification Prototyping is used for this purpose, there is no activity performed specifically to confirm the usability of the user interface. By contrast, this SSADM/GUI guide recommends:

- Modelling users and their tasks, environment, and usability requirements earlier in the development – as was done in *the 1st prototype simulation.*

- Developing an abstract definition of how the user thinks of the system (*Users' Conceptual Model*). This is comprehensible to the users, useful for their tasks, and is capable of being mapped onto the SSADM *Conceptual Model* (Logical Data Model, Entity-Event Models, etc).
- Developing this abstract view into the more concrete External Design products of Function Definitions and UI Design.
- Prototyping and evaluating the initial UI Design with real end-users, and iteratively redesigning it until the UI Design satisfy usability requirements.

It is this user-centred design approach with its core structure but adapted to meet the specific project needs which in this case blends with rapid prototyping techniques – so-called 'extreme programming'. The principle of the evolving university thus allows for the training of staff with frequently updated training material whilst retaining basic components of the interface that staff readily recognise, as identified by the pilot study results. This complements the approach toward the development of adaptive hypermedia that can accommodate such changes without rewriting the whole structure since the underlying conceptual model remains unchanged.

6.16 Implications for the wider issue of simulation in surgery

The vocational training of anyone whose work involves clinically invasive procedures requires the best from that individual with regards to decision-making and psychomotor skills, thus the development and monitoring of these is vital. Ultimately simulator-based skills training will become integrated within the framework of clinical governance (the *quality assurance* of clinical service delivery). The methodology of the *1st prototype simulation* task analysis for this trial was based upon the practical limits of what can be achieved with current available equipment, and also the expert analysis of tasks (61;139). The results of the *1st prototype simulation* trials were inconclusive with a small number of trainees, as would be anticipated in a hypothesis-generating study with a small population. A larger number of trial candidates are required. This led to the Virtual University supported approach in the VOU3 trials.

The tutorial system needed refinement as well as the simulation system, and so parallel development of the different domains (library, virtual classroom, discussion groups and the dynamic review process) meant that by *the VOU3* the user population and the tutorials were supported by a fundamentally different infrastructure than the original EVOU1 shoulder arthroscopy tutorial system developed for *The 1st prototype simulation.*

6.17 Conclusions based upon the 1st prototype simulation trial

The results of this EVW Pilot study suggest that such interfaces are to be built into educational environments to assist the control of training simulators. They include the development of the Virtual Orthopaedic European University (96) (voeu.ecs.soton.ac.uk) and the continued development of the shoulder simulation system into latest prototype. Both of these are expanded upon in later chapters. Ultimately such controls can be developed into standards in their own right. The aim of future work is to ensure that these devices assure compatibility between the intra-operative tools developed for the training and operating environments, and both objective skills tests and the subjective comments of the trainees are a promising start for the integrated system concept.

7 *Virtual university infrastructure facilities*

7.1 Introduction

The *virtual university is a monolithic infrastructure that provides tools for clinicians supporting the learning of clinical skills and information.* The web services are XML-based and so the view one sees is dependent upon the user's personal profile. This takes into consideration the training grade of the individual, their location and the deanery responsible for providing their training. The principal concept is to provide individuals with their specified learning material. It is structured as *essential and important further reading,* based upon their level of knowledge, state of preparation for examinations, attendance at specific courses, monitoring of clinical posts etc. New material from audits and trials can be integrated seamlessly.

The workload of individuals is negotiated and agreed as part of structured learning agreements that are managed on-line. It is tied in automatically to the clinical post, mapping user experience to the syllabus and the curriculum. The system, having been designed for surgery, has successfully undergone formal usability testing by clinicians but is generic and can be adopted by all medical disciplines and is now ready for *implementation and continued development* with dedicated partners, adapting the technologies to the specific needs of user groups.

7.2 Educational process design

It is clear that all aspects of simulation are managed in parallel and the process requires significant investment in resources. For this reason the project effort took a necessary change of direction to accommodate the effort of obtaining support for this method of working. The *VOEU* project was born. This thesis is however oriented toward the development of the video world infrastructure, which is then integrated within and supported by the VOEU, which is not the main focus. The rest of this chapter therefore outlines the educational strategy developed in tandem with VOEU.

7.2.1 Engineering usability into products

Gould, J.D., and Lewis, C., in their article: 'Designing for *usability* – key principles and what designers think of them' referred to the principles for developing usable products (142). These included focusing early and continuously on users, creating functionality that is likely to be used, and making changes before they become too expensive. By making documentation and training easier to develop and reducing the need for updates and maintenance releases the system should improve efficiency over paper-based equivalents. Integrating consideration of all aspects of usability and testing versions with users early and continuously, are integral to iterating the design.

7.2.2 Defining the goals and concerns that are driving the test

Surgeons are busy individuals with varying degrees of background knowledge. Sources of information and professional goals thus reflect this diversity and concerns include the task analysis and quantitative usability goals. These define the general goal and also quantitative goals for the system. This includes ways to measure wrong choices and the time taken to make a decision. It is also necessary to define the general and specific concerns for the test. There are also timely issues, such as staff training responsibilities. *Version 4*, which will stem from the work of this thesis, would benefit from improving this further, either by heuristic analysis or an expert review. With respect to the system itself, different groups of tasks from previous tests of this or other systems may be incorporated, but the heart of a 'popular' test is brevity.

7.2.3 Who needs a virtual university?

A virtual university infrastructure should be accessible by all potential users of University facilities. To implement this approach, authors of material must select the appropriate categories in the metadata accordingly. This background, blended with the variety of user experience that is encountered, necessitates virtual infrastructures for the training and evaluation of surgical performance, requiring reliable robust systems for the collection of data from both real-life and simulated environments. The data can then be used for both teaching of novice surgeons and the comparison of effective training techniques for different types of surgical intervention. The approach adopted involves providing a toolkit, which assists the process of consolidating data from different domains.

7.3 Tools developed in VOEU

7.3.1 Provision of customized multimedia educational modules to the user

VOEU is based upon the most recent technologies and standards for Virtual Learning Environment (VLE) development (143;144) to accommodate these interrelated clinical, educational, and research experiences. The repository for educational packages is referred to as the Multimedia Educational Orthopaedic Modules (MEOM) component of VOEU, built as an XML application based on Information Management System (IMS) Metadata and Content Packaging standards (145).

Using IMS nomenclature, every educational item is treated as a resource whether it is a whole tutorial, a web page, multiple-choice question (MCQ) or an image. Each resource is described by Learning Object Meta-data (LOM) and can be used and reused in different contexts as the educational content developer chooses. In the current version of VOEU the content can be organized in a varied but static way *i.e.* it is based upon hierarchical structures of its governing content-packaging *Document Type Definition (DTD).* In later stages, a transition to schemas was adopted to increase interoperability and compliance. The aim is to develop 'learner tracking' allowing dynamic content-building according to the individual's profile, similar to the earlier Intelligent Tutoring Systems (ITS). The MEOM development concentrates on two areas: building an automated orthopaedic knowledge repository and serving it to the learner in the most educationally effective way.

Every educational resource, accompanied by its metadata, will form an XML instance/document validated against the *VOEU metadata.dtd* and stored in the repository currently holding over 500 documents. Content providers will be able to add resources automatically through an *MS Word™ .doc* form as a template with a metadata header. Conversion of HTML to XML is handled by *'in house'* software developed by the IAM group in the University of Southampton. The intention is to provide customised multimedia educational modules to the user.

7.3.2 Library facilities in VOEU

Multimedia Educational Modules provide the declarative (factual) base of material for the education of the users. Reference material is structured for individuals based upon their user profile. This includes text still images, video and simulation in some cases. See Figure 7.1.

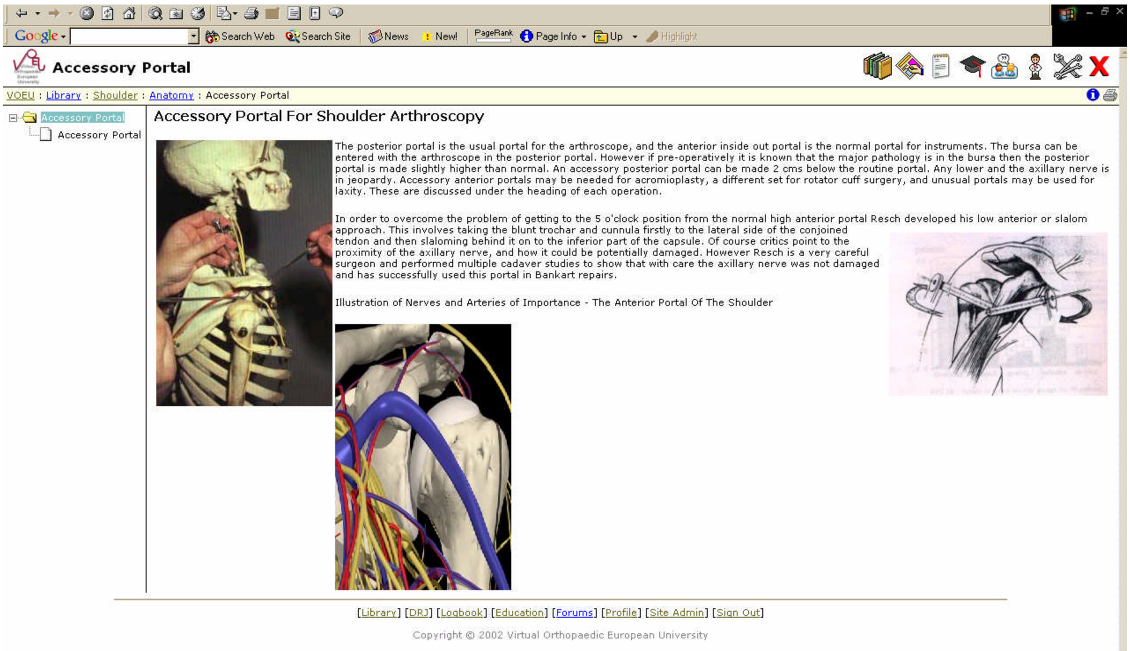


Figure 7.1 A screen shot of Library facilities in VOEU. Reference material structured for individuals based upon their user profile. This includes text still images, video and simulation in some cases.

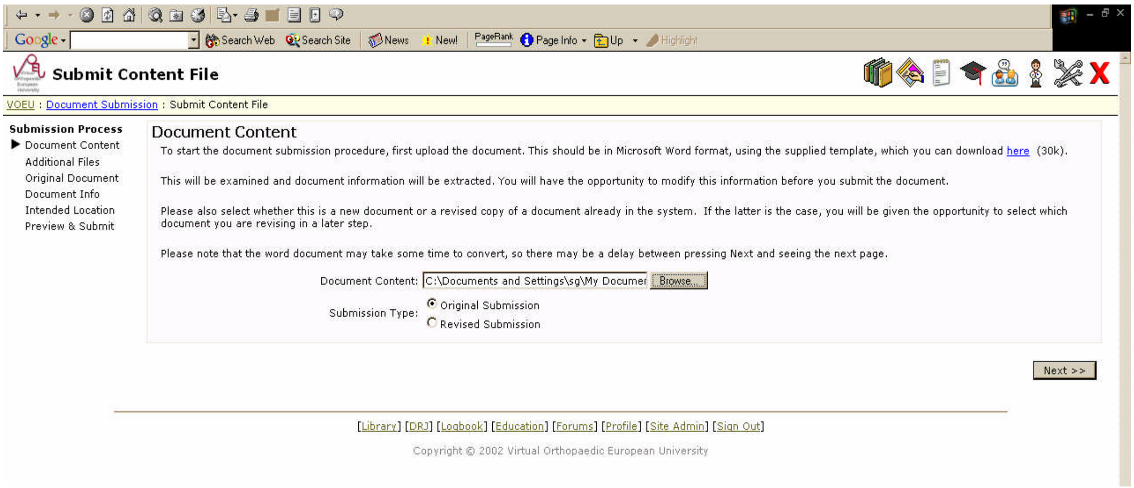


Figure 7.2 Document content is selected and uploaded using ftp from a local disk

The current VOU3 allows for the following:

- **Peer reviewed** structured content.
- **User uploading** and reviewing of documents using a simple *Microsoft Word*™ template.
- Provision of a process of **Non repudiation.**

- **'ICT for healthcare professionals'** available on-line and as a hardcopy for users.
- Support for **‘Critically Appraised Topics’** development of Evidence-Based Medicine.

Proposed functions for development include **full multimedia** user centred authoring environment. Political effort to establish **user groups** for the continuing updating and generation of new material is essential to ensure continued use. The system is already IMS and ‘Dublin core’ compliant.

There is a need for a peer-reviewed uploading process as demonstrated by the following screenshots. This is ostensibly mirrored in the review process, as each aspect of the uploading process is peer-reviewed. To assist navigation through the process for the user, there is a stepwise menu guide to the left of the screen. Over 500 documents have been uploaded using this method. The four steps demonstrated are:

1. **Document content** is selected and uploaded using ftp from a local disk (Figure 7.2).
2. **Document information** is checked against the metadata uploaded in the Word™ template to ensure it represents the appropriate 'Dublin Core' information and is selected for the appropriate intended audience, either by the author or later by the reviewer (Figure 7.3). This includes the author’s perceived audience, targeting specific groups of users.
3. The Intended location of the file within the author’s subject hierarchy is selected. This view may later be changed by the superimposition of ontology, such as a structured surgical course model (Figure 7.4).
4. The **document content** is then reviewed to ensure it is as intended (Figure 7.5).

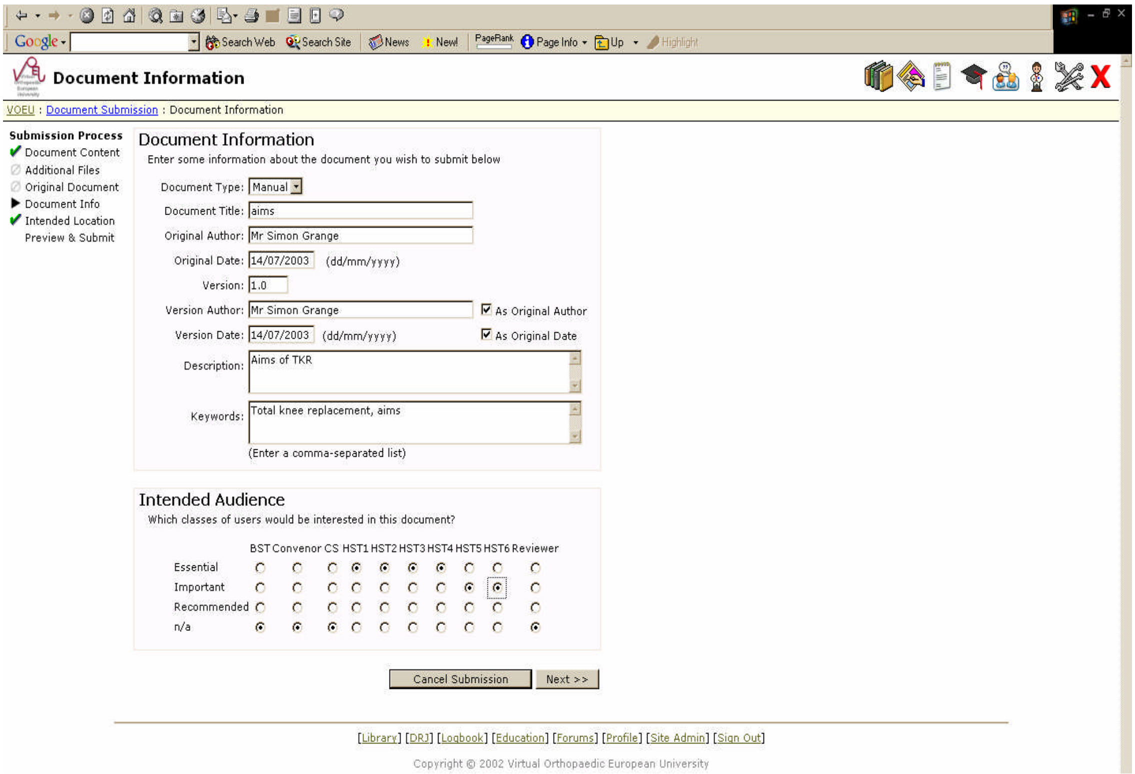


Figure 7.3 Document information (Dublin core compliant)

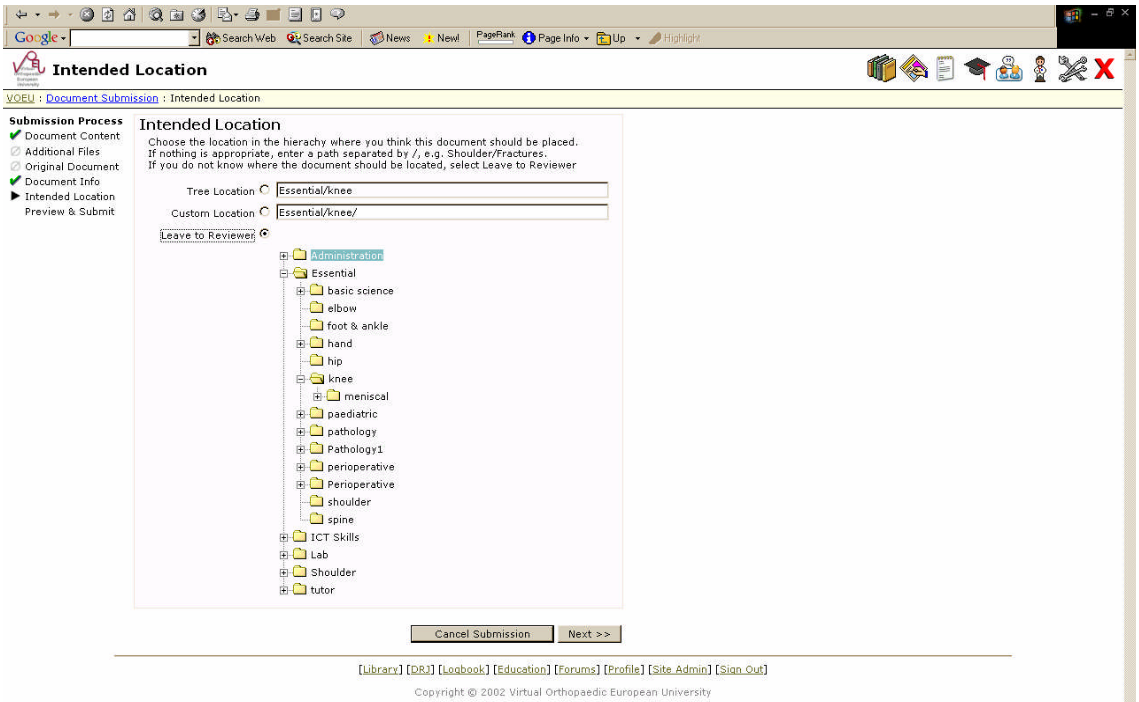


Figure 7.4 The intended location of the file within the authors subject hierarchy is selected. This view may later be ontology-driven

7.3.3 *Dynamic Review Journal facilities in VOEU*

This is an on-line shared working environment for managing trials both for audit and research. This service also provides a framework for collective contribution to data collection and paper writing. This creates a Virtual Observatory for the collection of data from simulation systems and the actual intra-operative data collection - see Figures 7.6 and 7.7. This developed into the Dynamic Review Journal (DRJ). Whilst all tools are relevant, this one is expanded upon in more detail as it forms the key interface for intraoperative data collection. This tool manages the clinical evaluation, adopting a similar presentation style to the Static Review Journals (*e.g. Journal of Bone and Joint Surgery*).

It is designed to meet the user requirements for a dedicated 'journal' of future surgical education systems. This should allow progressive case analysis and seamless integration with archiving. There will be case data mining that is constantly updated supporting search functions for the records. Unfortunately, demonstrating the benefit of differing clinical approaches upon a small population of cases without a considerable variation in outcome is not suitable for statistical analysis. For this reason, scalability was demonstrated using demographic study data of over 2800 cases.

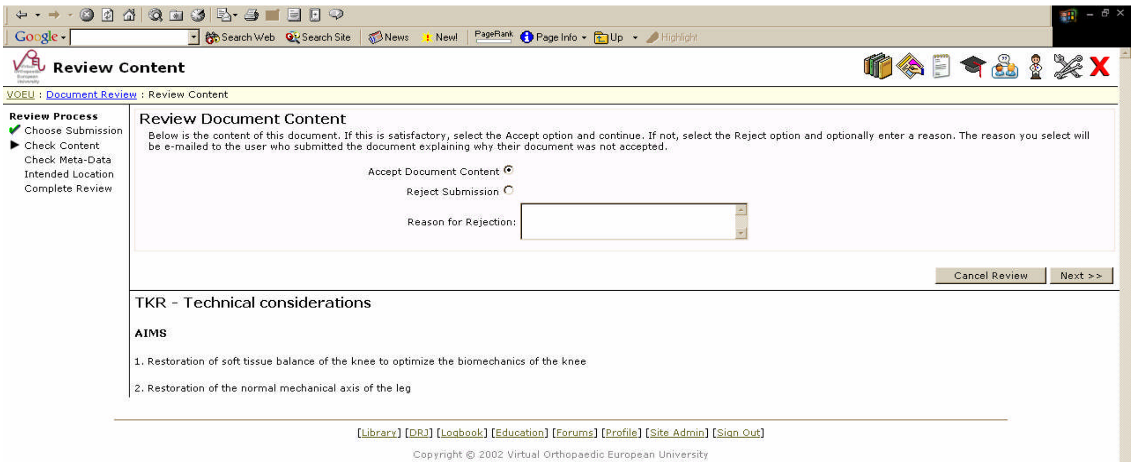


Figure 7.5 The document content is then reviewed to ensure it is viewable, as intended.

This demonstrated that the statistical engine performed adequately. Technical and ergonomic evaluation to detect technical weaknesses in current Image Guided Orthopaedic Surgery (IGOS) systems can build upon this.

To assure quality, the cases that are output by the DRJ must be technically and medically validated. The technical validation is managed by constraints on data collection. The medical validation is by the peer-review of case results on an annual/progressive basis. This process provides internal peer-review by the authors and selected panel, but relies upon an independent peer-review process before the document is uploaded in order to be accessible to others. A Static Review Journal (SRJ), which both stores and archives preprints and published articles (termed *ePrints),* is currently available. Results are only credible if they are peer reviewed and are available for scrutiny. The DRJ provides a record for non-repudiation of results. An infrastructure has been developed to support this by logging an individual's performance in conjunction with the development of multimedia tools to guide the trainee, see Figure 7.1.

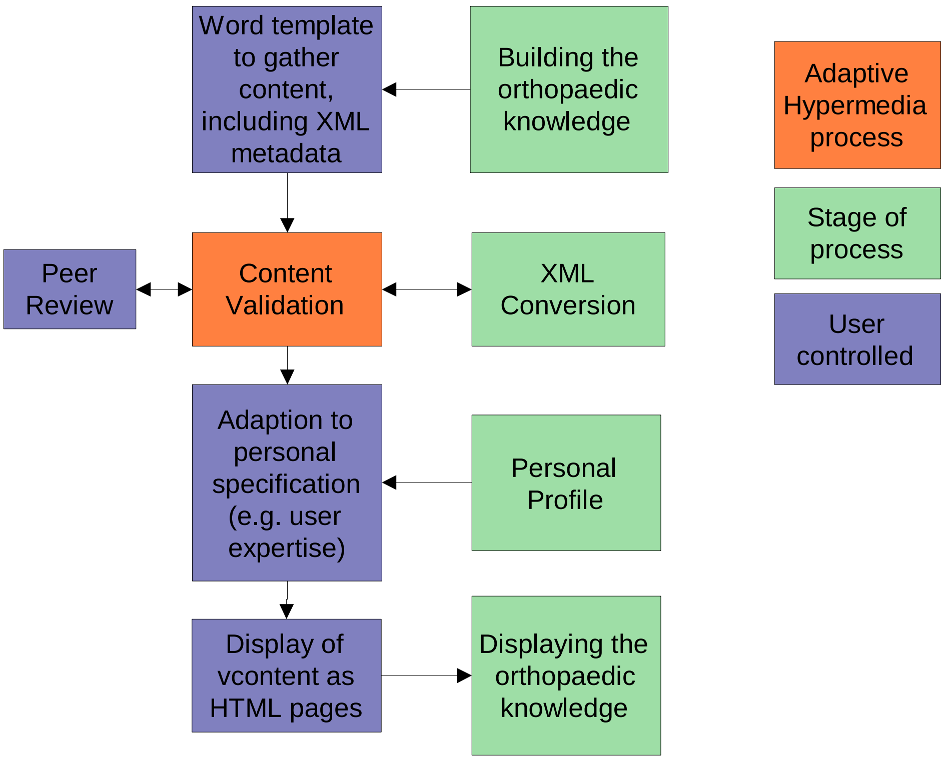


Figure 7.6 Document uploading procedure - HTML communication with XML forms

This will be the next step for simulation data collection via the DRJ. This could be employed intra-operatively for trainee tracking when using navigation systems. It is an important part of the process of transition to evidence-based medicine. The current version supports On-line trial template generation, web form trial data entry, and use of an on-line statistics package, multiple Journal formats and an alerting system. Dynamically reviewed material can be submitted to the library peer-review process to ensure up-to-date material based upon established clinical trials and on-going trials in the case of case reports and alerts.

The proposed functions for development include: enhanced trial data entry and heuristic support for users unfamiliar with statistics management. Journal submission templates for all leading journals (upon request) will be included, also offering a vertical prototype exploring other document templates e.g. case reports and editorials. A database should be developed that will provide a foundation for future research and audit. Finally it is envisaged that there will need to be forwarding to national and international trials centres for analysis.

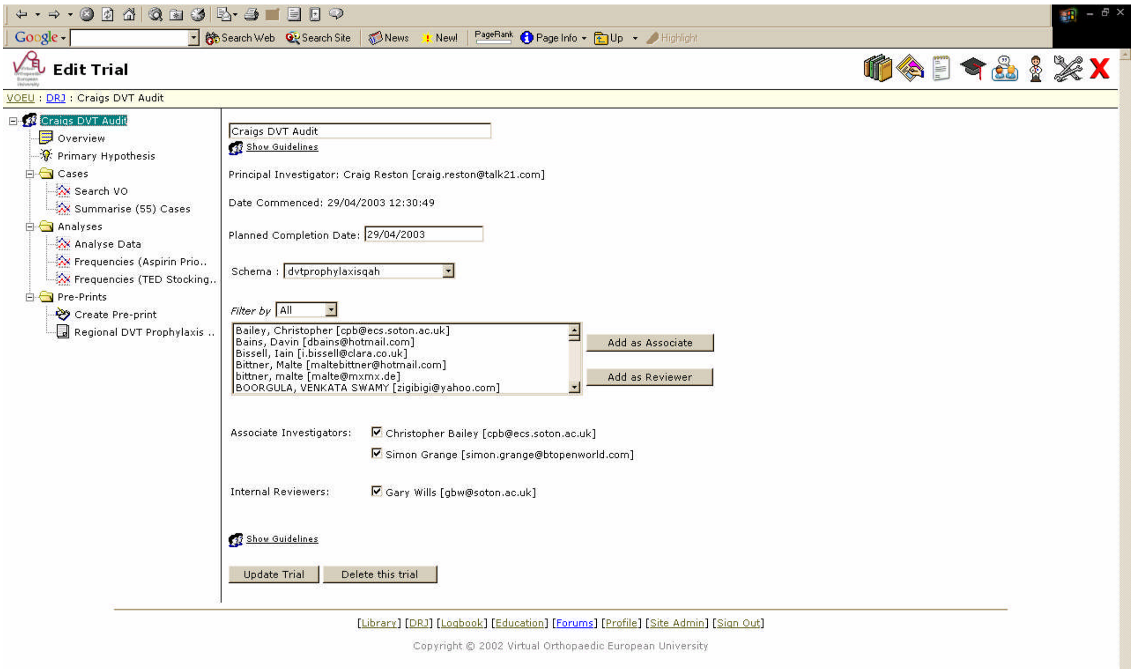


Figure 7.7 A screen shot of Library facilities in VOEU. On-line shared working environment for managing trials both for audit and research. Embedded simulation

Embedded novel modalities of simulation (36;65) have been tested for the emulation of surgical procedures for training and experimentation focusing upon image guided surgery - see Figure 7.8 demonstrating version 2.

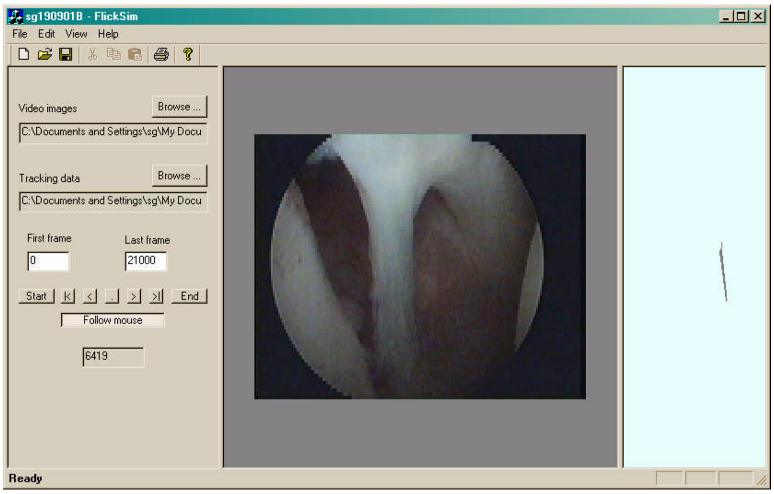


Figure 7.8 Embedded simulation (for shoulder arthroscopy in the trials)

7.3.4 Surgical Logbook facilities in VOEU

The logbook module offers the ability to log clinical cases (36;65). This is essential for the recording of surgical procedures for training and management of trials, see Figure 7.9. The current version is BOA compliant with the User profile linked so that users have a Personal case base linking to the dynamic review journal for audit and case review that links to their participation in trials. The intention is for the system to be able to accommodate Regional In-Service Training Agreement (RITA) Summary Sheets and to offer automatic uploading of data to other records, such as the BOA logbook JCHST, GMC revalidation etc. All of this will be facilitated by pervasive access from handheld devices.

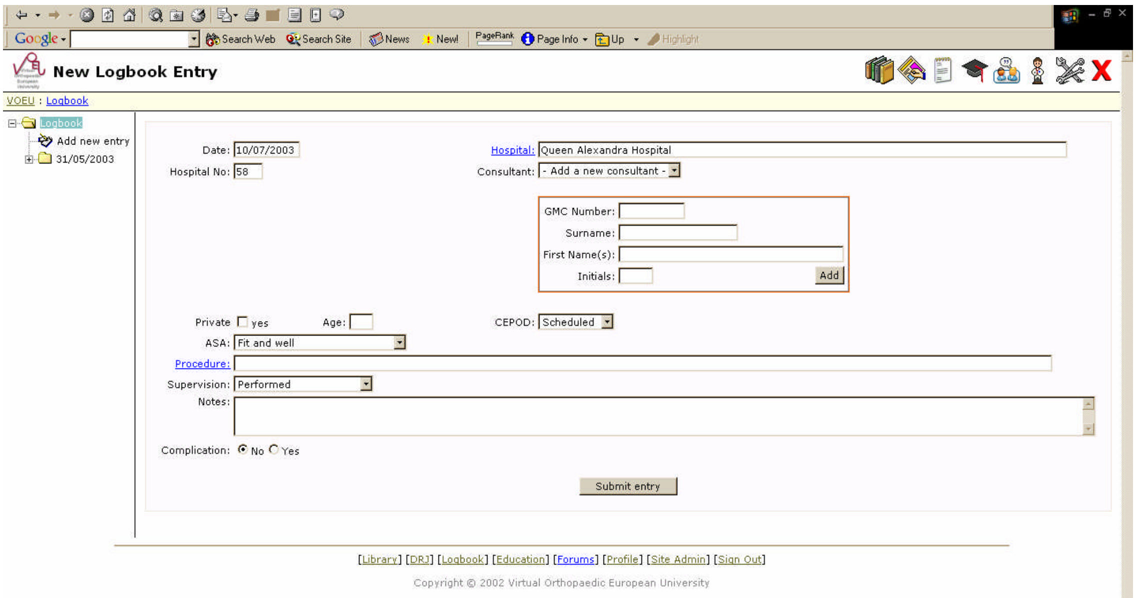


Figure 7.9 A screen shot of surgical logbook facilities in VOEU. Conventional surgical case record with 'drop down' menus of cases and surgical team members to facilitate data entry

The integration of procedural skills monitoring into the user educational record (the surgical logbook) is an integral part of the dynamic record of surgical performance and is currently a topic of debate since trainees are obliged to use a commercial package that has been imposed upon orthopaedics in the UK. There are however certain basic standards that can be adhered to. Data relating to cases according to the basic requirements of Higher Surgical Trainee's logbook (see Table 7.1, Figure 7.10) should be collected, including the following data recorded[29]. There are many other subheadings that can be added, but these will be attached to these groups. Cases will be filtered according to the types of Case / Operation (see Figure 7.10), whether the patient is Adult / Paediatric (up to 16 years old) and whether the operation was assisted, performed supervised or performed unsupervised. A summary page based upon the above classification is provided allowing for expansion to accommodate the additional features being evaluated (type of equipment used) will be included.

Hospital Initials	Hospital Number	Date	Patient's Initials	Age	Sex	Operation	Comments Specifications	Involvement	Complications	Consultants Initials

Table 7.1 Basic surgical logbook headings

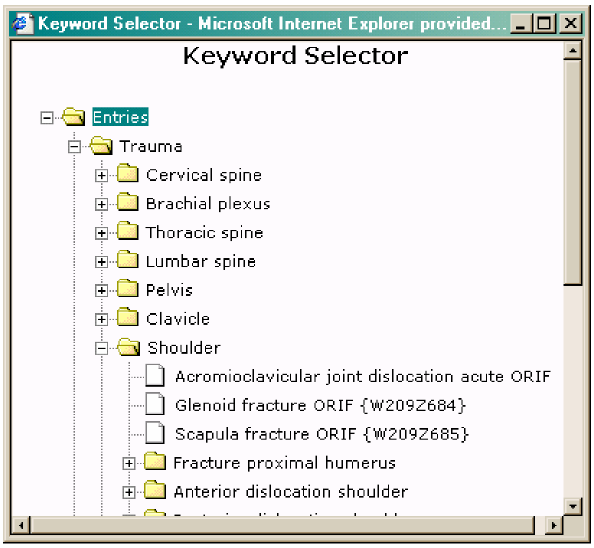


Figure 7.10 Context dependent key word selection for logbook cases

[29] Confidentiality and Security must be maintained in accordance with the Data Protection Act – and the user should be registered voluntarily with the Information commission. www.dataprotection.gov.uk/

7.3.5 *Virtual classroom facilities in VOEU*

Case based learning is supported with automatic linking to cases of appropriate selection for user expertise, and tied in to the learning resources in the library. A catalogue of cases is available to the user based upon their experience (see Figure 7.11). In the current version this includes context-dependent keyword searching, and the ability to construct new courses from established material with images and links to video possible in the cases. The proposed functions for development are:

- Support for the educational and training responsibilities of the department.
- Facility for users to prepare on-line educational material, using a case-based structured approach.
- The feedback loop for user performance in tests
- Development of feedback based upon written trainee responses

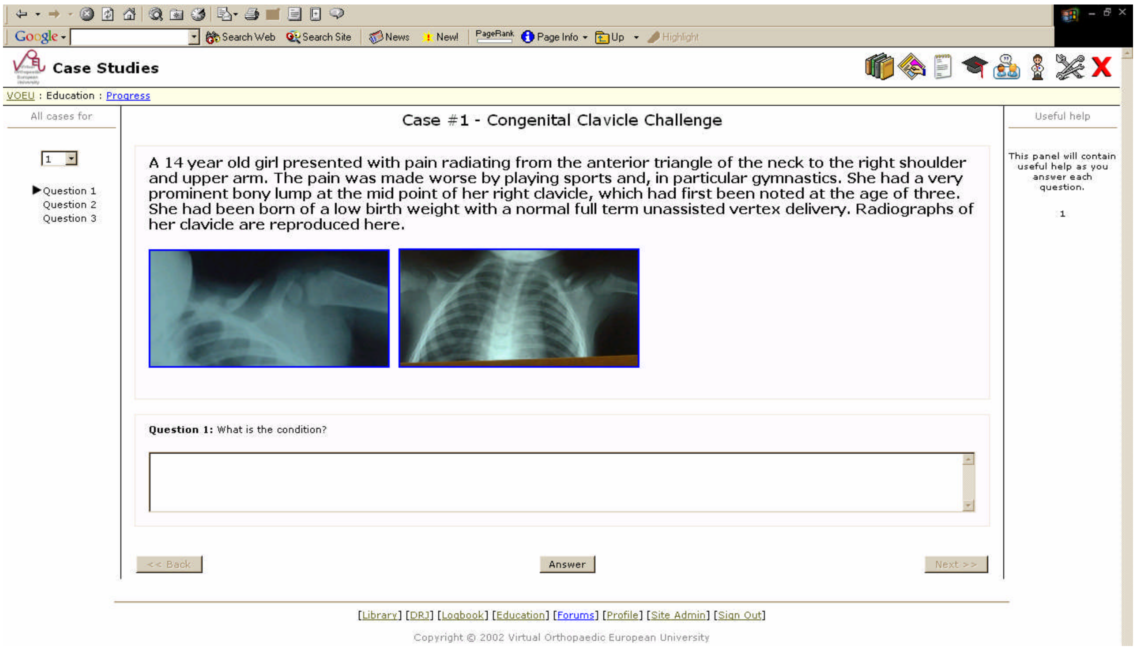


Figure 7.11 A screen shot of virtual classroom facilities in VOEU

7.3.6 *Discussion fora facilities in VOEU*

This is a shared workspace that is secure for members of different discussion fora. The moderator regulates communication. Users can thread the discussions and target specific topics, see Figure 7.12. In the current version the functions include:

- Lists of active forum participation
- Ability for users to create new fora.
- User roles as participants and moderators.
- Automatic private discussion forum generated for paper authors linking to their draft articles.

The intention is for this to evolve into themed discussion fora with issue tracking and user monitoring by agents of broader discussions.

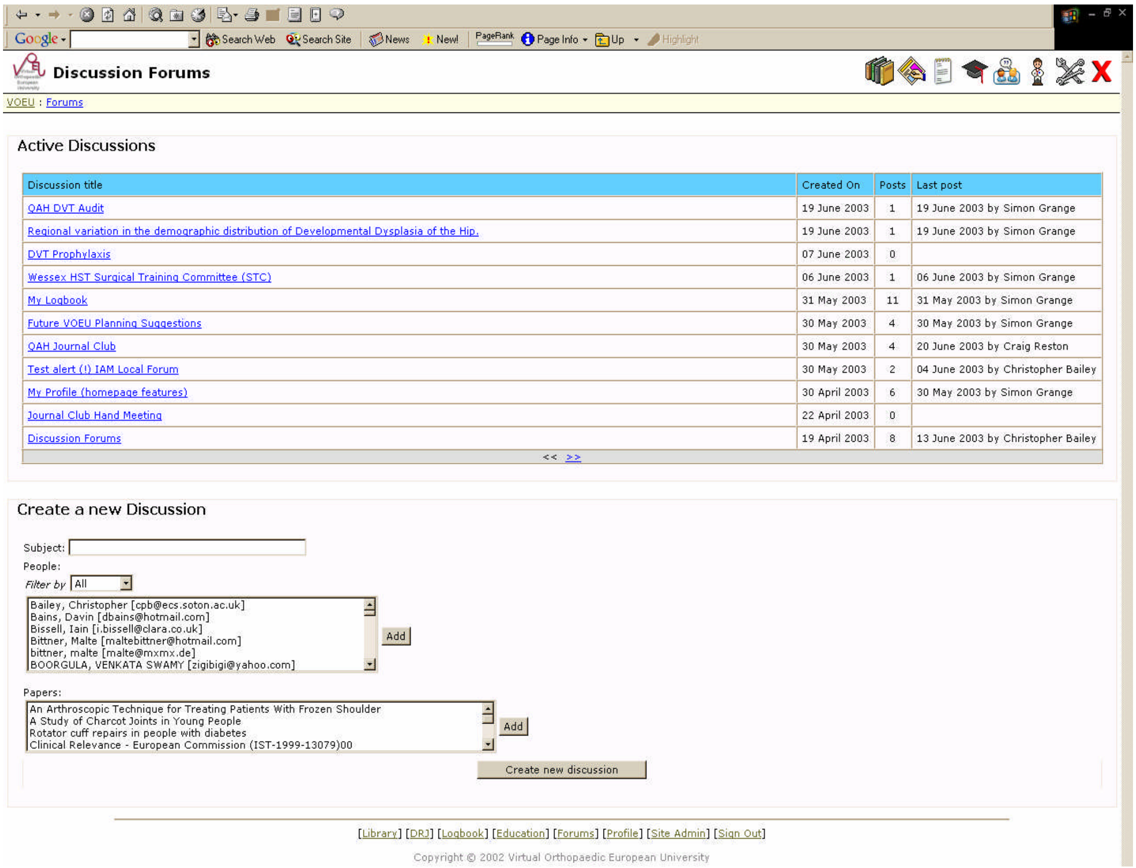


Figure 7.12 A screen shot of discussion fora facilities in VOEU

7.3.7 Personal profile facilities in VOEU

This is where the user develops a personal view of their postgraduate training and continuing professional development including individual learning agreements. All the involved staff groups (stakeholders) should be able to access the system to include and obtain the information commensurate with their duties (see Figure 7.13).

In the current version it is possible for the user to track Case progress. Users can update their password, subscribe to courses, initiate learning agreements and maintain their profile. Their Learning Agreements are negotiable with the tutor, considering curriculum, syllabus and clinical post issues. There is dependency on Feedback forms to improve the service. To improve the service further it is proposed to build the forms into the e-portfolio in the form of a ‘Dynamic’ CV, providing support for CV generation based upon experience and activities. This will offer access to the journals and databases that they have previously organised permissions for by adapting the roving profile inside the hospital. By building upon the plans for pervasive access mentioned above, intelligent agent monitoring of activities will assist users to access available resources in conjunction with other organisations. This may be extended to support study-leave administration for hospital staff in accordance with the hospital requirements.

Combining the above tools within *one working environment*, the virtual university infrastructure (113) aims to meet the needs of orthopaedic surgeons to combine clinical, educational and research responsibilities via a platform accessible over the *internet*. VOEU was specifically established to address the issues of data access, presentation and development of a structured learning environment for the training of novice surgeons, which it has achieved, though a key by-product of this is simulation integration and access to the operating theatre.

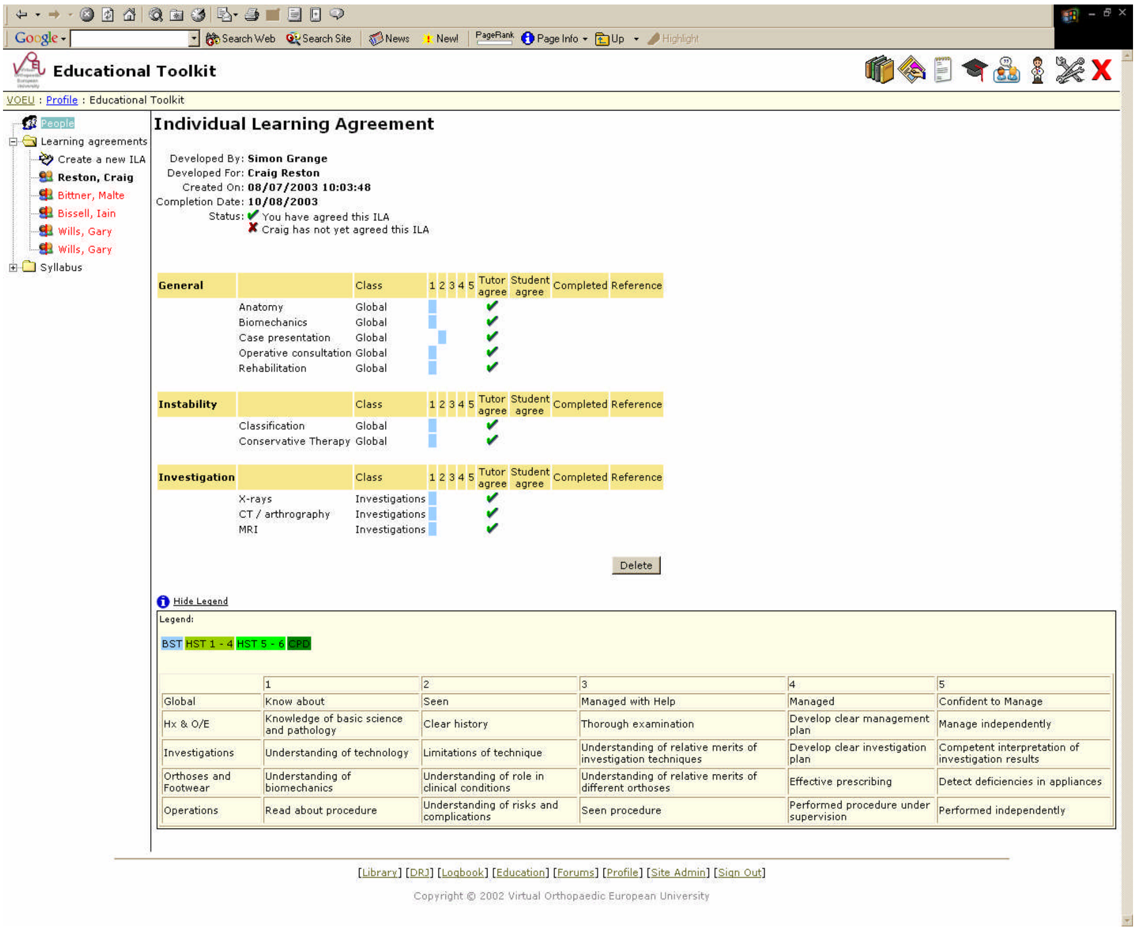


Figure 7.13 A screen shot of personal profile facilities in VOEU. The User profile links in their clinical experience with their educational programme

7.3.8 Administration facilities in VOEU

This section of the service is available to the administrators to facilitate user education, forum management, notices and user records, see Figure 7.14. Whilst this currently supports an alert (Notice board) system that all users can employ to target specific groups of individuals instantly along with Library, Trials Template and User Management, the intention is for this to expand also, with the following functions:

- Pervasive access to resources and tools.
- Delivery of services using a modular approach.
- Prioritisation of tasks in accordance with user demand to ensure primary tasks are achieved in time and secondary tasks are completed in order of priority.
- Appropriate organisational labelling and design identity.

7.4 Trainers in teaching surgery - the educational contract

For healthcare professionals, much of the workload has always been directly involved with education in various forms, targeting various groups of trainees. The use of ICT provides us with an opportunity to improve the efficiency of both teaching and learning in the context of Life-long learning (CME / CPD) (146;147).

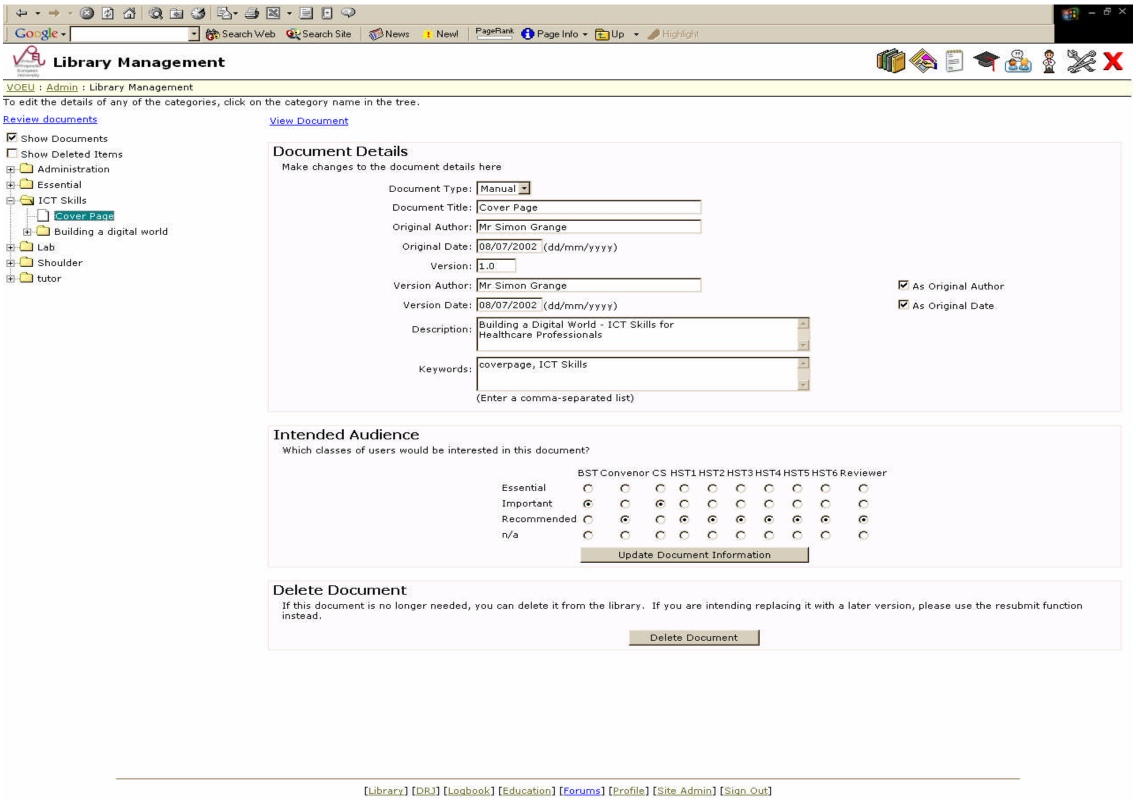


Figure 7.14 A screen shot of administration facilities in VOEU

With the introduction of virtual infrastructures, it is possible to manage most of the administrative, research and educational workload of the '*university*' within the digital domain. This has potentially huge benefits for the users and other stakeholders with greater access to information, thereby reducing the friction associated with traditional education and communication infrastructures. It is the ease of cost-effective access and also the standardisation of languages and development of stakeholder-friendly authoring tools that is precipitating a '*paradigm shift*', with the world rapidly segregating into those who have access and those who do not.

Educating users so that they understand and accept the cultural change will be of far more value (human capital) than technologies in the long term. This takes longer to develop and is harder to specify. One approach is a planned contractual model for trainees and trainers that forms the core of the educational contract. This will extend the role of the training institution as the management of clinical placements becomes more apparent in this process. The templates for the structuring of learning agreements based upon the syllabus play a key role here and one such solution is demonstrated in Figures 7.15- 7.17 below.

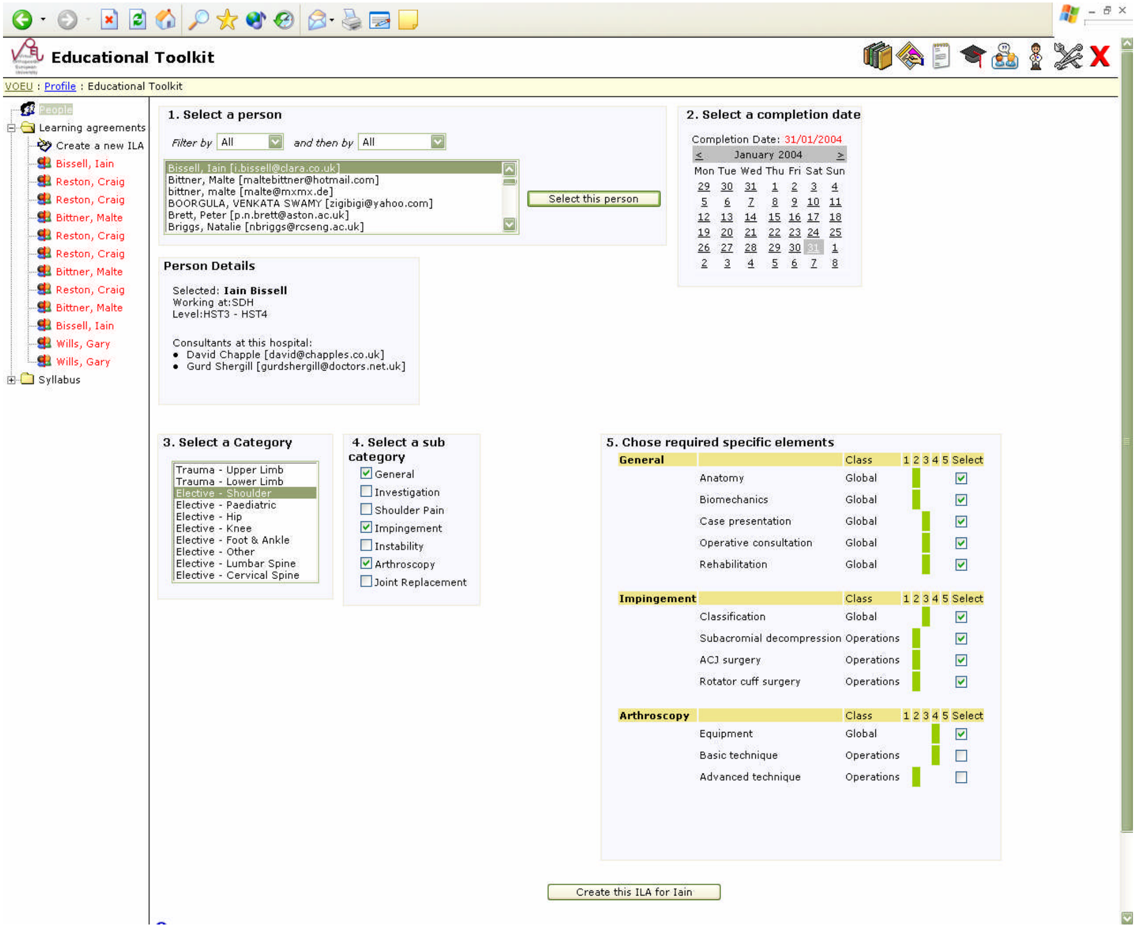


Figure 7.15 - Selecting relevant parts of the syllabus

The process involves selecting relevant parts of the syllabus using the on-line course infrastructure, negotiating the contract using learning and performance objectives and selecting relevant parts of the syllabus using the on-line course infrastructure. Part of the educational strategy employed relies upon the evolving university concept to allow for the updating and upgrading of educational material in light of new results, using analysis of incoming data from ongoing clinical trials for the evidence base. This acknowledges that a virtual university is a living infrastructure that evolves with time, due to changes in both its underlying philosophy and its staff. Since surgeons are mobile well-educated individuals whose work demands excellent availability of educational material that is up to date and focused on their particular learning situation, the approach must reflect the need for life-long learning material to be collected as well as the 'Just In Case' archives. New material is being collected constantly and this needs to be properly evaluated and integrated into the learning infrastructure appropriately. The educational contact between the trainer and the trainee is designed to structure this process, so helping to construct the surgeon's personal ontology.

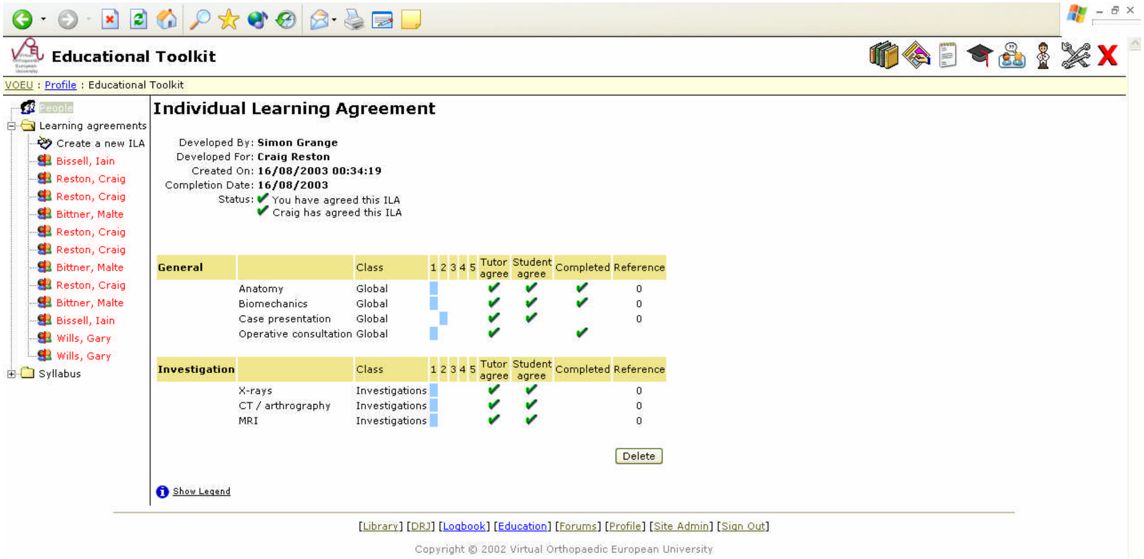


Figure 7.16 - Negotiating the contract using learning and performance objectives

7.5 Summary

The tools and resources, *learning objects*, will be flexible enough to enable their use at a number of levels, from major pedagogical re-engineering of courses by convenors to looking up isolated topics. The process of using the environment consists of the following stages:

1. Mapping of curricula to pedagogical approaches.
2. Identification of appropriate teaching and learning methods.
3. Evaluation and selection of appropriate resources.
4. Identification and integration of resources and tools.
5. Creation of examples.
6. Delivery, evaluation, and refinement.

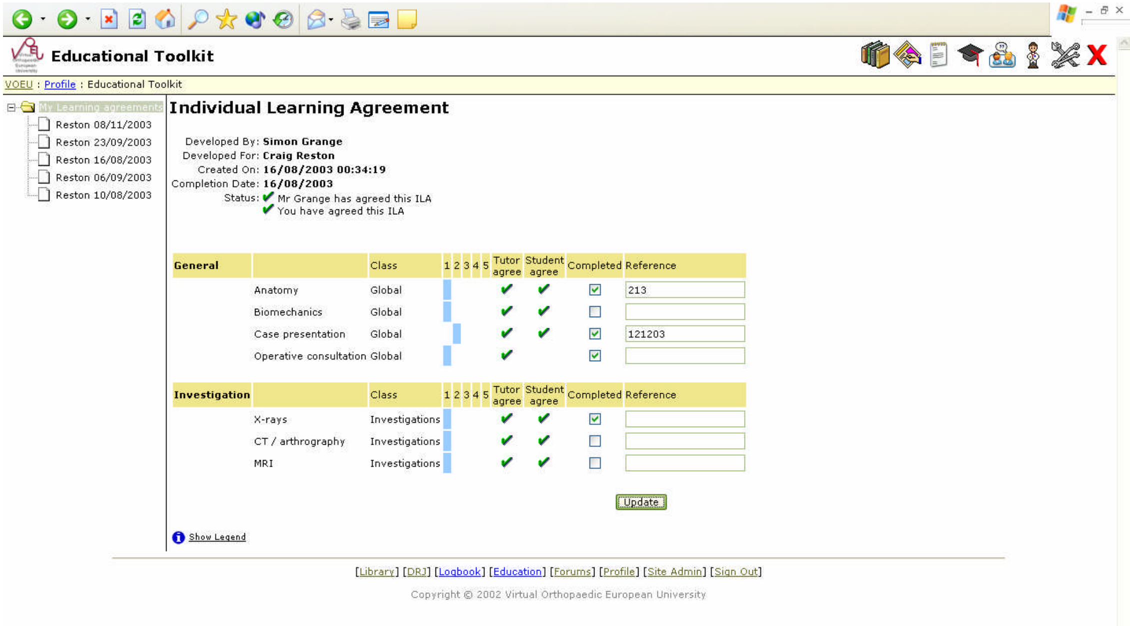


Figure 7.17 - Monitoring an individual's progress through completion of tasks

The associated research questions focus on the following themes:

- **Pedagogical** – What new pedagogical models are possible with innovative uses of new technologies and what is their impact, Will this approach ultimately lead to transformations in practice as intended in the *3rd prototype simulation?*
- **Socio-cultural** – What organisational and human dimensions and fundamental changes in practices are emerging? What changes (organisational, strategic, and cultural) must occur in order to formulate future implementation plans?
- **Technical** – How interoperable will the different systems and architectures be? There is a political directive to merge the technologies with other systems. By applying learning object metadata (IEEE metadata guidelines), user-generated learning materials can be developed from distributed resources.
- **Legislative** – Different IPR issues concerning ownership need to be resolved. The system needs to meet changing legal requirements and national agendas. Seamless training from medical school to the coaching of others should be well served by these technologies.

8 *An evaluation of the virtual university infrastructure as an effective learning environment*

8.1 Introduction

Since web-based training is becoming the dominant method of flexible distance learning and presents a practical solution for the collection of performance data from disparate sources, earlier work (the version 1 simulation) demonstrated a need for a cross-platform tool to facilitate such data entry and analysis. This was crafted in the 2nd and 3rd prototypes. These main experiments suggest the thesis, which asserts that a structured on-line learning environment offers an alternative to the use of textbooks providing focused educational material in the most appropriate situations.

The Version 2 (EVW 'Pre-study') trial was performed to confirm the structure of the VOU3 (main study), outlined below. Version 2 study was a large focus group established to provide the initial feedback upon the initial virtual university model integration and to obtain a broader range of opinions from the planned end-user group. Version 2 provided feedback for prototype 3 and the more sophisticated final study, and also its proposed methods for data collection. Learning the lessons from this led to the next iterative design cycle and the VOU3 evaluation thereof.

VOU3 offers both better integration and evaluation. Importantly this was achieved whilst testing a geographically isolated group of surgical trainees. This is described and commented upon. Obtaining automated results from a web-based training tool suggests the need for integration of databases and the need for standardisation of assessment criteria, and also their organisation using a data repository. Development of a large enough population to compare performance necessitates different authorities submitting XML filtered non-uniform data, which raises issues of compatibility and the need for design criteria that can be imposed to facilitate this process.

8.2 VOU3 simulation trials validation methodology

As stated above, web-based training services are likely to become both the dominant method of flexible distance learning and performance data collection from disparate sources. The cross-platform tool was tested using the VOEU toolkit to support this analysis, specifically the DRJ. This is protocol-driven, detailed in Appendix 8.1. The web-based platforms tested in *the 1st prototype simulation* included versions of Internet Explorer and Netscape Navigator. Future competition between the main players in the field is likely to impact upon the design, with the need for distributed database access having a significant influence upon the potential for XML implementation. A design strategy that minimises this risk would be of benefit. *MS.net* and *MS Internet Explorer 5.5* or later, which includes *XML 1.0* support, was therefore adopted. Usability engineering involves identifying the user groups, analysing tasks and setting usability specifications. It leads to developing prototypes, which are then tested via iterative cycles. This approach is incorporated into *the 3rd prototype simulation*.

8.2.1 Strategy for version 2 simulation trials (pre-studies)

A cluster of small pre-studies was performed to confirm the structure of the main study. This study cluster consisted of two main parts:

1. A focus group was established to provide the initial feedback on the *initial virtual university model* integration to obtain a broader range of opinions from the planned end-user group. This was a very low yield (12 out of 72 participants completed their questionnaires) relating to the earlier simulation model (Figure 7.4) and so necessitated the changes that led to the simulation in the *3rd prototype* and also its evaluation.

2. A pilot study for the usability analysis of VOEU was performed to assess the effectiveness of the VOEU system. This was satisfactory and so the *Version 2* form has been incorporated into the main study accessed by users as *'Questionnaire A'*.

8.2.2 Results of the version 2 simulation trials

The 'pre-study' provided feedback for the design of the final study and its proposed methods for data collection. This emphasized the need for greater integration of end-users into the design process. It was clear that mass demonstration of the system with users observing and then completing questionnaires did not adequately represent the processes involved, and so the yield of completed questionnaires was inadequate (only 12 of N=72 *i.e.* 17%). The role of evaluation includes the collection of users' opinions and experiments to confirm these findings, as in *the 1st prototype simulation.*

To monitor the overall performance of the simulator, it is necessary to gain adequate data on the performance of a population of individual surgeons. This should be large enough to provide an adequate distribution curve for elucidation of the *'auto-testing'* functions and also to be able to compare the system to traditional teaching methods. Video can offer other attributes such as registration of information, if this is recorded in association with the specific frame locations. This can later be used for labelling or assessment of an individual's performance and was beyond the scope of the *Version 2* trials.

Demographic records are used to subcategorise the operators by virtue of their experience. The intention is that, as a greater number of records are obtained, it may be possible to use the simulator in a distinguishing role, based upon user experience, to differentiate groups of surgeons that require further training or show particular merit.

The learning curve is long and the significance of this is that it should be possible to obtain a clear threshold that an individual must reach before he or she can practise the techniques in vivo. As a database of performance expands, more specific data may be extracted, but this will only be effective if 'like' qualities are compared. An example of trends that could be modelled is performance vs. the following variables: years in service, years in surgery, age and hand dominance. The gender and training grade of the surgeon were also recorded.

Some of these factors may raise issues that could conflict with the wishes of a patient or the socio-political climate of the time. This cannot ultimately be ignored, but the interpretation of such data must be taken in context. Such simulation systems can test only individual subsystems and not all the abilities of a surgeon. One must always consider psycho-social skills, knowledge base, and a myriad of other factors in conjunction with an individual's results of skills cluster monitoring.

By logging an individual performance, monitoring their track through a joint, and comparing this to the envelope of acceptable performance as defined by an expert arthroscopist, one can assess accuracy (a psychomotor skill) as well as time to complete a task. Although this is not a critical factor, it is an indicator of the familiarity of a surgeon with the particular operating environment.

To ensure fair assessment the simulator must collect the spatial performance of an appropriately matched peer population. This is another *auto-testing* function. The database that develops will continue to reinforce the distribution curve of group performance, the *caveat* being that if the simulator is modified, the performance being measured may change. Hence the distribution curve may change as indicated by the AIPES scoring system. This suggests that non-parametric testing would be more appropriate. If this is found to be a normative distribution of data then it may be more efficient to use the Dynamic Review Journal (DRJ) to apply parametric tests for data with a normal distribution and then cross check with non-parametric tests. The aim is to correlate the two and if they correlate well, the parametric tests would be used, as they are more efficient, requiring fewer subjects in the study *i.e.* a higher power.

Once a successful range of performance is defined, it is reasonable to assess an individual performance and provide a quantitative peer-reviewed value, the significance of which may still be debated for the reasons mentioned above. Then one could compare the performance of different simulators, assuming that if one has a large enough population of practising surgeons, the populations will be comparable. Ideally, one would be able to use similar scenarios to compare the populations in different geographical locations to confirm this. Certainly, using a remote access networked system for simulation would facilitate this.

8.3 The VOU3 simulation usability test management

The purpose is to define the goals and concerns that are driving the test, to decide who should be participants, and organising participant recruitment. It is necessary to select and organise the tasks to test by creating task scenarios, and in so doing decide how to measure usability. Preparing other material for the test, and preparing the test environment, also includes preparing the test team – assigning roles and training team members. It was essential to supervise the practice before the start by conducting a pilot test, making changes as needed. This consisted of a small group (N=8) that was then included in the main trial results.

The usability engineering approach with previous testing and the iterative design strategy provided experience as to deciding who should be participants, developing the ‘user profiles’ and considering relevant characteristics that all users have. Foremost is computer experience. Since the technologies used centre around a distributed architecture the author considered that the users should be broadly grouped into the three categories: no experience, no network experience and network only experience[30]. This does not exclude the possibility of other user experience. Deciding which factors matter most when considering the surgeon's work experience is easily ascertained from their job description. So when recruiting participants, and selecting and organising tasks to test, the profile of the users becomes important to gauge the ability of the system to meet the user expectations and abilities. An example of this is in the creation of task scenarios for learning.

Another issue is deciding how to measure usability. Due to the user population characteristics, it was necessary to be able to swiftly manage the interviews with minimum disruption to normal routine. Recruiting a manager for this and also suitable training for them to organise the interviews were essential. This role included preparing the test environment, preparing the test team assigning roles, training team members and practising. The role of trials manager was delegated to Dr Craig Reston.

The prototype 2 'pre-study' provided feedback for the design of the final study and its proposed methods for data collection. One issue in prototype 3 is the need to evaluate both the content and the delivery system together since prototype 2 has in effect managed the two separately. For this reason the issues surrounding the educational infrastructure are addressed first, followed by usability issues. The Virtual Environment Simulation Interface used in prototype 2 developed with Dr Chris Setchell offered greater interactivity. It was a stand-alone application and could not be easily developed further in this format. It is illustrated above in Figure 7.8.

8.3.1 Decisions regarding the VOU3 simulation trial management methodology based upon version 2 results

Employing the VOEU system in *the 3rd prototype simulation* allowed for adoption of the flexible data collection system. This has been described in more detail elsewhere (148). In the trials this is now managed as an *.aspx* form-based process with data collected in the *'My Data'* area directly. The results are presented in the format of the initial trial and represent a collection of data that was collated both digitally and by paper forms.

[30] This is so future training needs to ensure equal entry level skills can be met by such courses as the European Computer Driving Licence. www.ecdl.co.uk

8.3.2 Testing the educational infrastructure

The approach adopted employs checklists for end user evaluation of the system. The protocol for the trial is provided in Appendix 8.1. The checklists employed in the clinical trials are derived from the work of D. Pascal Girard (UJF) and Dr Gary Wills (University of Southampton) relating actual practice to four categories being evaluated:

- Relating tasks to the pedagogy
- Assessment methodology
- Internal structure and content
- Multimedia Educational Orthopaedic Module

To obtain suitable candidates for the data collection a registration process selected staff from another of the UK training regions (Wessex). The process for selecting for trials management and data entry is expanded upon in the following Table 8.1 and detailed in sections 8.4 to 8.9. This may still induce a selection bias. The selection process is detailed in Figure 8.1.

	Section of study	Questionnaire
A	Introduction	First time users are registered for demographics
B	Demographic	(1) User Clinical Experience + Usability (General) (2) User ICT Literacy
C	Familiarization with the simulation system	This involved a 'hands-on' supervised session
D	Simulation Task Performance	Technical feedback for systems administrators; setup efforts (planned future development for Multi-centre trials in *Version 4*)
	User simulation perceptions	AIPES simulation assessment
E	Test Performance	User Questions in the Virtual Classroom regarding Educational Effectiveness to evaluate the educational effectiveness of the entire multimedia package, since the AIPES form only concentrates upon the simulations.

Table 8.1 Outline of study questionnaires (templates)

8.3.3 The VOU3 simulation testing strategy

This represents the simulation validation and verification process outlined in Chapter 4 in Figure 4.1. The strategy relies upon five subsections labelled A – E handled using the questionnaires as discussed above. The testing strategy outlines the duration of the test (about 45 minutes) and the individual author's requirements to prepare his or her own content with references where the relevant material is available. This is described in the protocol for data submission (anonymisation). The principles of non-repudiation and peer reviewed authorship and also intellectual property protection by *Copyright Authorisation* as with conventional journal articles, are also apparent.

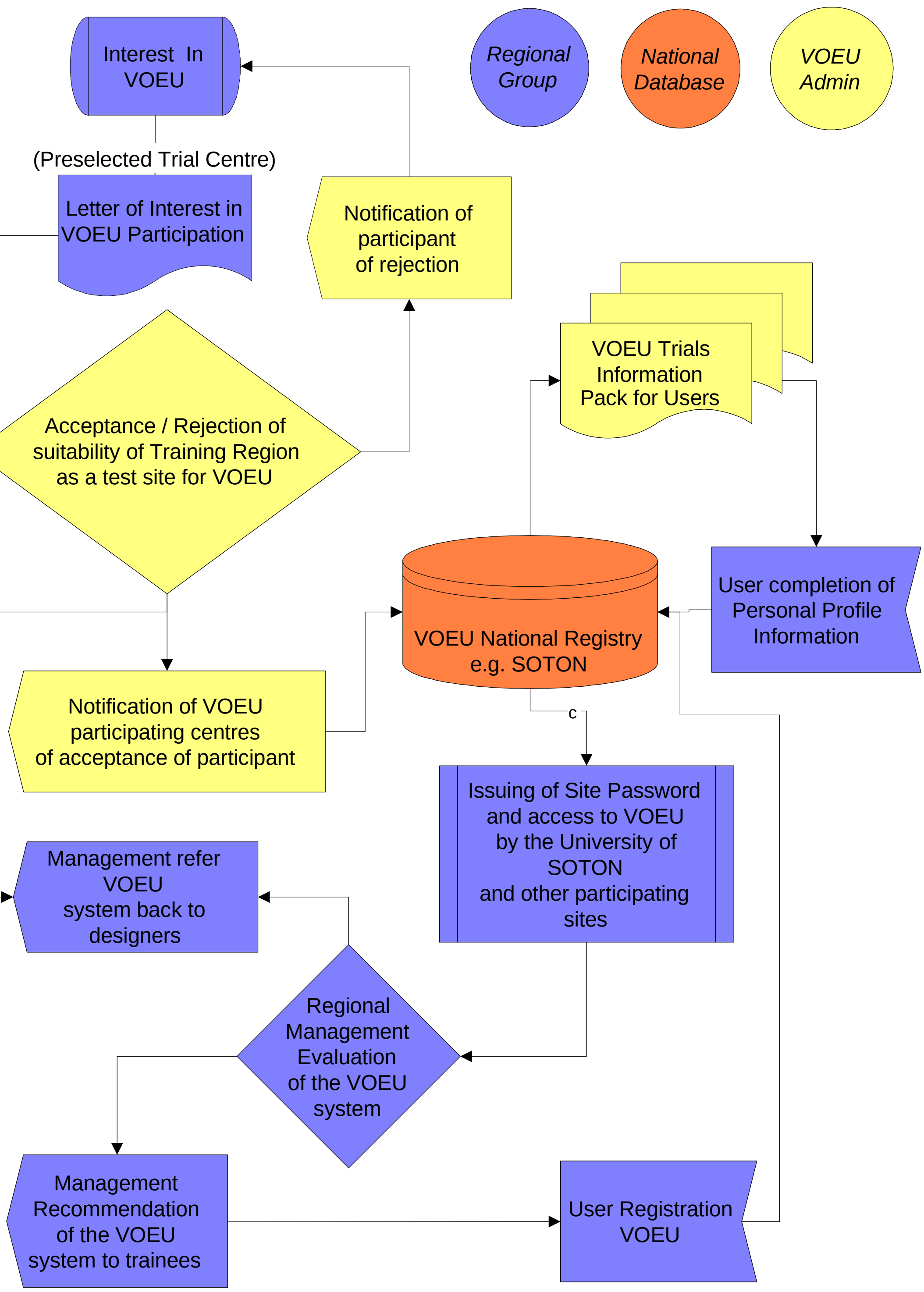


Figure 8-1 Plan for participation in a VOEU trial such as the VOU3 simulation trial

8.4 A – Introduction to *VOU3* trials results - user profiling

The results from the trial of the web-based training tool, testing groups of surgical trainees, are described and commented upon in this section. The principle is one of continued analysis of the results that can be reported upon, involving obtaining automated results exploring the integration of databases and the need for standardisation of assessment criteria and their organisation using a data repository. Development of a large enough population to compare performance would necessitate different authorities providing non-uniform data, which raises issues of compatibility, and design criteria that should be introduced to ease this process.

8.5 B - Demographic - usability analysis of surgical interfaces

It is necessary to draw distinctions between *Functionality analysis* and *Usability analysis.* Usability is an attribute of every product, so what is usability? The aim is effective (quick and easy) accomplishment of the users' own tasks *i.e.*, goal-directed behaviour. Ultimately Human–Computer Interaction is only human interaction via a computer. Human beings interact via the modalities of our senses whether directly or indirectly, using computer networks.

Focusing upon user-centred design means usability focuses upon users and for the study it includes both potential and actual users. For example the users' managers cannot substitute for users. Dumas (149) argues that people use products to be productive. Such tools should therefore demonstrate:

- Ease based upon effort (time, number of steps, skill level) required using a system for what they want.
- Success in predicting an appropriate action to take.
- Interface and documentation vs. performance goals (effort *vs.* expectation).

Users need to know the goals and tasks clearly. Users are 'busy people' trying to accomplish tasks; tools need to facilitate, not hinder this. Hardware and software are tools to achieve results, so usability should increase productivity. Documentation is simply a tool to increase the use of hardware and software. Users have a low tolerance for time taken learning and using tools.

Dumas states: *"Users decide when a product is easy to use."* This means that there is a balance of the following factors influencing the time and effort input vs. potential benefit. The factors relating to this are:

- Time and effort that users are prepared to spend learning
- Performance changes, through the effects of usability.
- User Interface issues: Windows, Icons, Menus, Pointers
- The adequacy of information sources such as the accompanying 'Manual', 'Quick Reference Guide' and 'On-line help'.

These are not separate entities. Users need consistent documentation. An iterative design approach to usability engineering has been built into the development process. Users were actively involved in the development process, allowing users' needs to drive design decisions. Work was carried out in a team with skilled usability specialists (*e.g.* Dr Gary Wills), with interface designers and technical communicators involved. Quantitative usability goals were set early in the design process, and testing of the products for usability *e.g.* the VOU3 prototype, was considered integral to the design cycle. By building usability testing into the design commitment, this demonstrates that the technology is working.

8.5.1 Approach to testing future surgical educational system

Providing an Educational Strategy is one aspect of the development of course infrastructures. The course convenors need to have a degree of freedom as to how they structure their courses for the following reasons.

- Not all subjects lend themselves to a rigid teaching methodology.
- The course material has a very high turnover rate in some cases as new material is published frequently.
- Making the course curricula models (educational course models) provides the convenor with the options that can then provide feedback – via results as to which are more or less useful components.
- Adopting such a system allows the status quo to be maintained whilst new courses are developed and adopted, so as not to disenfranchise those course convenors who are currently computer-illiterate.
- The **Orthopaedic Computing** course held in 2000 was an initial effort by the RCS in association with the European Commission to address this last issue and to foster Surgeon-Engineer relations inside the VOEU project framework.
- The recruitment of users is based upon selecting those with the necessary skills and the user has the option to remain anonymous.

8.6 B – Demographic - results of the prototype 3 analysis

Initial alpha testing by the author is inadequate as it introduces bias, since the author is familiar with the system and not a realistic end-user. Field trials for observational evaluation include the following strengths:

- Users are the actual *intended* users, well matched to the target group. They are of varying levels of experience, which are taken into consideration.
- Field-based studies avoid bias introduced by the laboratory, but suffer from distraction, ambient noise, and interruptions. The context is set, which is particularly important for the simulation.
- Data can be collected in a non-invasive manner using *'.aspx'* forms completed at the end of demonstration sessions.
- Evaluation reflects the iterative design process that will incorporate the results of the next generation of simulation design.
- By using the DRJ the results can be updated as more data becomes available.

The observational trials can be used to support the process of engineering toward the target of integrated simulation (as demonstrated in the 1st prototype simulation). Such trials compare the generations of design and the process of gaining a clearer insight into the real world of the end-user though they are not adequate to set standards of conformance. In effect the process of simulation analysis outlined in the VOU3 for the AIPES scoring system has its roots in observational analysis, and represents an experiment for testing of the system against user expectation. The purpose of the simulation analysis is thus twofold: (a) to ensure the method of evaluation for such systems is adequate to relate to the user's requirements, and (b) to evaluate the conformance of the simulation system to the AIPES protostandard. One can deduce which areas are to be targeted in future research and development. This relates the user's response to established scoring systems novel to the application. To attempt to validate this by comparison, the NASA TLS scale was adopted. This has been tried and tested for evaluation of tasks for simulation of space missions. The blend of knowledge-based and procedural skills required lends itself to this adaptation.

8.6.1 B1 - Results of user clinical experience

Following an introduction (Section A of the interview) aimed at settling candidates into the trial environment, taking approximately five minutes, the candidates were asked to complete the demographic questionnaire (Section B of the interview). This first addressed the question of user experience, both of surgery and of computers. The following demographic information was collected prior to evaluating the system using the questionnaire. Appendix Table 8.1 focuses upon general experience of orthopaedics (years in specialty) and specific experience/expertise in the area being assessed (shoulder arthroscopy).

Analysis of the user clinical experience results raises issues such as *confidence* reflecting the *value* of clinical experience. All individuals of a higher grade (more experience) returned confidence scores higher than more junior staff performing the procedures. This is the same for all studies in the *trials* series.

- **User hand dominance:** was not retested in the main trial, as adequate results are available from the pilot study. It was considered that this should not affect the process of equipment design.
- **User status:** Users were asked their name and email address and could optionally provide information regarding their acquired status in surgical training (N=13).

The findings displayed in Appendix Table 8.2 suggest the population is a reasonable reflection of the prospective users for the clinical educational tools, which are aimed at both basic and higher surgical trainees.

- **Age and years of service:** The user profile included their age (Appendix Table 8.3) at the time of the trial. Younger users may be more aware of the potential routes for using digital technologies. The average age is 31 years (Appendix Table 8.4) with approximately 5 years of service experience within surgery.

The findings suggest the population is a reasonable reflection of the prospective users for the clinical educational tools, with a wide range of experience. This compares reasonably favourably to 5 years 4 months (N=16) in the 1st prototype *simulation* pilot study.

- **Use and attitudes of networked digital technologies:** The combined findings (Appendix Table 8.5 & 8.6) suggest the population is a representative reflection of the prospective users for the clinical educational tools, who are progressively using networked technologies as part of their normal daily lives and work practice (N=19). The trend is toward greater usage both inside the health service and for personal use. The combined findings (Appendix Table 8.7 & 8.8) suggest the population of the prospective users are aware of the potential values of such technologies, based upon their previous experience.
- **User experience of shoulder arthroscopy:** This is part of the user profile but was included in the questionnaire upon surgical simulation to match the experience to the procedures being validated. The questions were changed in part from the previous pilot

study questions that considered confidence with the procedure in order to reflect the exposure to the surgery in more detail. The reason for this is that the likely Individual Learning Agreement (ILA) design will include the logbook input that will map the cases the surgeon has experience of, in accordance with the surgeon status (*i.e.* performed or assisted etc.) and so this was applied here. Appendix Table 8.9 demonstrates the 'Number of Arthroscopies' performed during the pilot study as a comparison. Data collected allows us to consider the VOU3 population in a little more detail (N=13), at a layer of granularity below the layer above.

- **Analysis of the results of user arthroscopy experience**: The findings in Appendix Table 8.10 suggest the population is a representative reflection of the prospective users for the clinical educational tools, with the caveat that the population is small and the opportunity for exposure to operative training is diminishing. It does suggest that the data should be collected in accordance with the regulations of the British Orthopaedic Association (in line with the *Joint Committee upon Higher Surgical Training (JCHST)*) so that the level of involvement with each case is recorded. This is part of a process being developed inside VOEU to take the logbook of the user into consideration and integrating this with individual learning agreements to provide the foundation for a competency base.

8.6.2 Conclusions drawn regarding the user profile

The population is a reasonable reflection of the prospective users for the clinical educational tools. One user provided the following subjective comment with respect to the control of the trial: *'I need more practice'*.

This reflects the fact that the system was being evaluated as part of a beta-test *'in the field'* using real users in their actual place of work, offering the advantage of realistic setting but emphasising a fundamental issue for surgical populations- the demand upon their time. They are understaffed in the UK and so training opportunities for trials such as this are inadequate. There are however differences between the two populations of the users that reflects the change in demographics between the populations with the time frame in which the 1st prototype simulation and the VOU3 trials were set. This also reflects the potential reduction in training opportunities that has arisen due to the changes in junior doctors' training. This has concerned some clinicians as it is likely that the same volume of training opportunities may no longer arise, due to the European Working Time Directive ('maximum hours worked') requirements. The longer-term impact is yet to be appreciated. Teasing out these two factors will be necessary and requires more data hopefully provided by the DRJ in due course. With the exception of one German Senior House Officer (initially trained in the DDR) all doctors in the second study have less clinical experience than the comparative group evaluated initially.

8.6.3 B2 – Results of user ICT experience

- **User ICT literacy -** This was modified from the earlier experiment to consolidate the questions to a core block that was representative of the user population's requirements. Candidates were only obliged to answer questions concerning the applications that they considered relevant to their work (N=19). The reader needs to consider these results in light of the pilot study results (not a direct comparison) showing some improvement in general ICT literacy. The VOU3 results are focused upon the skills actually required to use the virtual university system.
- Analysis of user ICT literacy results are addressed in light of the plans for the evolving Virtual Orthopaedic European University (VOEU) and the general improvement of ICT skills around the world. Whilst 1% of the world's population own a computer, in the earlier study at least 90% of the test population did. Less than 1 in 20 have had formal training upon how to use it for their work. Between 1 in 10 and 1 in 5 have attended specific courses for applications. This led to the main study protocol development including the demographic review of whether users were active network contributors, such as the use of the *File transfer protocol 'FTP'*.

When considering the VOEU system, users demonstrate familiarity with the use of web-based systems, using the web from home daily in 58% (11) of users, weekly in 26% of users and only one user (5%) using it around once per month and two users (10%) never using it from home. This lags slightly behind the 14 (74%) who use it daily and 4 (21%) who use it weekly at work, with only 1 (5%) using the web rarely at work. When asked if the users found the high-tech route to information intimidating, 4 (21%) agreed, 2 (10.5%) were indifferent and 13 (68%) disagreed. 8 users (42%) do and 6 users (32%) do not use file transfer protocol with the others uncertain.

8.7 C – Familiarization with the simulation system

14 (74%) strongly agreed that they would want to use the system on a regular basis whilst 3 (16%) were indifferent, one (5%) disagreed and one (5%) strongly disagreed. 6 users (32%) agreed that a paper-based system may be easier to use, whilst 5

(26%) were indifferent and 8 (42%) disagreed. 3 users (16%) agreed that a paper-based system was a more effective means of sharing information, though 4 (21%) were indifferent, and 12 (63%) disagreed with this argument. Certainly 16 (84%) considered that this is not the only method available for the retrieval of information, and 3 (16%) neither agreed nor disagreed.

There was positive feedback upon the use of the VOEU technologies for convenience with 16 (84%) agreeing, 2 (10%) indifferent and one (5%) disagreeing. 15 (79%) agreed that searching was easier, with 3 (16%) users equivocal, and one (5%) disagreeing.

Approximately the same response was given, with 15 (79%) agreeing and 4 (21%) indifferent to the statement that the system sped up the retrieval of information. 15 (79%) agreed, 3 (16%) users were equivocal, and one (5%) disagreed with the statement that the system was easier to use and that they would prefer to use it.

4 users (21%) strongly agreed that they found the system awkward to use, balanced by 4 (21%) others who strongly disagreed, whilst the majority 11 (58%) disagreed. 10 (53%) strongly agreed that they enjoyed working with the VOEU system, whilst 5 (26%) neither agreed nor disagreed, 3 disagreed, with one (5%) strongly disagreeing. Consequently 9 (47%) strongly disagreed that they would not recommend using the system, 6 (32%) disagreed, and 4 (21%) strongly agreed. Cancelling this ‘double negative’ in this statement it is interpreted that 15 (79%) would recommend the system to others.

8.8 D – Simulation task performance

This was tested and the results were described in Chapter 4. The AIPES evaluation forms the basis of the validation process for the embedded simulation.

8.9 E – Test performance - educational effectiveness

To collect the data required for the analyses referred to above, each VOEU member was introduced to the system in a 30-minute familiarisation session that exposed the user to the main functions of the site. They were then invited to explore the site themselves and to complete a few straightforward questions that reflected the main functions (applying appropriate functions to achieve set tasks), leading to an on-line question session. This included a question upon simulation usability.

The users were assured that the answers to this questionnaire and any additional comments would only be used as part of the VOEU project evaluation and associated academic reporting, such as this thesis. The data remains the property of the individual, held with their permission. By completing the questionnaire, it is accepted by the users that their results will be used for analysis and development of this system.

Each user's unique identifier was their email address. For the usability evaluation of the VOEU process the user population was small (N=19), but reflected the type of user group that would be likely to adopt VOEU. The system has been designed to support different user groups, stratified according to experience. One of the problems inherent in testing is achieving adequate numbers. The orthopaedic approach to this is a pragmatic one, comparable with the iterative design process. Users' time is precious, so accessing a population of testers at least provides results to work from and within the limits of this trial matching and stratification of the user population is secondary to achieving timely user impressions of the system.

Although the main study borrows heavily from the design methodology of earlier studies it is important to consider that the role of the questionnaire itself has changed. In the on-line form development of the later questionnaires, it has been possible to modify the sections to include the key subsections as separate entities. The structure of questionnaires outlined in the table below allows for modification of the sections according to planned future development by adopting a modular approach. Generic usability analysis forms are now presented with and without demographic data attached for the accommodation of first time users. This would allow for demographic data collection during enrolment.

This data was collected using an on-line form, submitted as 'Questionnaire C'. This can be located at: http://voeu.ecs.soton.ac.uk/library/doc.aspx?docis=428 and was prepared from an html draft with the assistance of Dr Chris Bailey. Details of this can be found in Appendix 8.2. The user review of the system included the characteristics of the VOU3 system. This included the following: impression, sense of command, navigability, ability to learn, helpfulness and perceived effectiveness. These are detailed below.

8.9.1 Command of the VOU3 simulation (VOEU) system

Six users (32%) strongly agreed that they were unsure if *they were using the right command.* However 10 (53%) disagreed and 3 (16%) strongly disagreed with this statement. 11 (58%) strongly agreed that they *were in control when using the system,* 3 (16%) disagreed and 5 (26%) were equivocal. Reassuringly, 16 (84%) strongly agreed *that the system was responsive to their inputs*, with 3 (16%) users not expressing an opinion. 7 (37%) strongly agreed that they found *the interaction with VOEU cumbersome,* though 8 (42%) disagreed, 3 (16%) strongly disagreed and one (5%) was indifferent. In total 11 (58%) did not find the system cumbersome. 17 (90%) felt the system *reacted quickly enough to their selections,* 2 (10%) did not express an opinion, and 16 (84%) strongly agreed that *it was easy to make VOEU do what they needed it to do*. 2 (10%) did not find this the case, and one (5%) did not express an opinion. Analysis of the *user's sense of command* of the VOU3 system suggests that VOEU demonstrates good controllability. A minority felt that *it was difficult to be sure that they were using the right command and were not feeling totally in control.* One individual found it *hard to do what he needed to do with the system*. His comment was: *'I admit to being a technophobe'.*

8.9.2 Navigability of the VOU3 simulation (VOEU) system

To the question, 'In VOEU there are too many steps required to get to the information I needed' whilst 7 (37%) strongly agreed, 12 (63%) disagreed. This topic may need to be addressed with better short cuts. Fortunately 16 (84%) strongly agreed that they *were able to move around the information in VOEU easily* with only 2 (10%) disagreeing and one (5%) user equivocal in their opinion. Supporting this aspect of the layout, 15 (79%) strongly agreed that *the side toolbar provided useful shortcuts* whilst two (10%) disagreed and two (10%) did not express a position. 15 (79%) strongly agreed that they *found the VOEU menus and content pages useful* whilst 2 (10%) users were indifferent and 2 (10%) disagreed. Unfortunately 6 (32%) strongly agreed that they *often become lost/disoriented when using VOEU,* though 2 (10%) were indifferent, 8 (42%) disagreed, and 3 (16%) strongly disagreed. The positive perspective is that 11 (58%) did not become lost/disoriented, but whether these users had better ICT skills is not clear, since 5 out of the six were either daily Internet users at work or home.

13 (68%) users stated that *there were plenty of ways to find the information they needed* whilst 4 (21%) disagreed, with 2 (10%) not expressing an opinion, and 15 (79%)

users felt *the actions associated with the options on the side toolbar were easily understood*, with 2 (10%) disagreeing, and 2 (10%) not expressing an opinion. 5 (26%) strongly agreed that *the information displayed was inconsistent*, though 9 (47%) disagreed and 5 (26%) strongly disagreed with this statement. Only 3 (16%) strongly agreed that *they did not understand the icon in the menus* (all daily Internet users) whilst 11 (58%) disagreed with another 5 (26%) disagreeing strongly suggesting a split in opinion.

6 (31.6%) strongly agreed that *the screen became cluttered and confusing* whilst 8 (42.1%) disagreed and 5 (26.3%) strongly disagreed, so the majority (13, *i.e.* 68.4%) considered the display was not cluttered or confusing.

Analysis of the user's sense of navigability of the VOU3 system findings suggests the population generally felt *able to navigate the VOEU system*, A minority felt that there are '*too many steps required to get to the information I needed'*, and became lost/disoriented when using VOEU. Importantly 3 users commented that *the screen became cluttered and confusing. One individual found it hard to find the information that he needed to do with the system.*

8.9.3 Ability to learn of the VOU3 simulation (VOEU) system

16 (84.2%) strongly agreed that *learning to use the system was easy*, the other three users did not express a decisive opinion, though 6 (31.6%) strongly agreed that they *did not have enough time to learn to use the VOEU system,* which is understandable as it was only approximately 30 minutes introduction. The other 13 (68%) users disagreed. The fact that 15 (79%) disagreed and 3 (16%) strongly disagreed that the system was difficult to learn to use and only one (5.%) strongly agreed with this statement support this. The author interprets this as an indication that, with an adequate training service, the vast majority would easily learn to use the system. Again the degree of introductory training is an issue when one considers that 8 (42%) strongly agreed that *it was difficult to learn more than the basic functions of the VOEU system,* whilst 9 (47%) disagreed and 2 (10%) strongly disagreed. 14 (74%) users strongly agreed that *enough guidance was given before using VOEU* whilst 3 (16%) were did not express an opinion, one (5%) disagreed and 1 (5%) strongly disagreed. With respect to the flexibility of the interface, the author is reassured that 16 (84%) strongly agreed that they *felt at ease trying different ways to get to the information I needed*, with only 2 (10%) disagreeing and one (5%) strongly disagreeing.

The analysis of the user's sense of learnability of the VOU3 (VOEU) system findings suggests the population found learning to use the system was easy, though one user considered that they *did not have enough time to learn, and so for them it was difficult to learn more than the basic functions of the VOEU system*. On the whole the results were again very positive, though the user did not feel at ease trying the different ways to get to the information needed. This supports the argument for providing structured training and the preparation of the ICT handbook to assist such individuals who may not have had the same advantage of exposure to digital technologies as others.

8.9.4 Analysis of the user's impression of helpfulness of the VOU3 simulation (VOEU) system

5 (26%) users strongly agreed that the *error messages were not easy to understand*, but 13 (68%) disagreed with another not expressing an opinion. The fact that there was not enough information on how to respond/proceed to error messages was strongly supported by 6 (32%) users whilst 12 (63%) disagreed and one (5%) did not express an opinion. The author suggests that there is room for improvement here, since 9 (47%) users strongly agreed that *the system was awkward to use if I wanted to do something out of the ordinary*, though the other 10 (53%) users disagreed. Certainly this raises the issue of constraints upon data entry. 14 (74%) users agreed that *the system messages were helpful when coping with an error* and only 4 (21%) disagreed with one (5%) strongly disagreeing, and reassuringly 13 (68%) users strongly agreed that they *understood and were able to act on the messages provided by this system*. 4 (21%) disagreed with this statement, with one (5%) strongly disagreeing.

16 (84%) users strongly agreed that *the system help files provided enough information to use the system*, with the other three users is in the trial disagreeing. The analysis of the user's sense of helpfulness of the VOU3 (VOEU) system suggests that the population felt that there *was not enough information on how to respond/proceed to error messages.* All the users understood and were able to act upon the messages provided by this system and (perhaps optimistically) agreed that *system help files provided enough information to use the system*. This however may not be a true representation of the whole system, rather the scenario generated for testing. The matter is being addressed by making the ICT book available in a searchable format by topic on-line.

8.9.5 User's impression of the effectiveness of the VOU3 simulation (VOEU) system

When asked if 'Using VOEU would NOT be of use to themselves personally in their job', only one (5%) user strongly agreed, whilst the 10 (53%) disagreed and 8 (42%) strongly disagreed, suggesting that there would be almost unanimous uptake of the system with adequate training. The acceptance may not be quite so definite since 3 (16%) users strongly agreed that *'Using VOEU would get in the way of the task that they were undertaking'* though 10 (53%) disagreed with this and 6 (32%) strongly disagreed. Certainly the cultural acceptance would be eased by having an effective way of translating from an individual's conventional work practice to the 'Brave new world' model demonstrated by VOEU. An example of this would be the ability to integrate data from other sources such as *MS Excel™* spreadsheets. An almost identical number of individuals found it difficult to obtain the information that they needed when using VOEU I, with 3 (16%) users strongly agreeing, whilst 13 (68%) disagreed, and 3 (15.8%) strongly disagreed.

Again there are a minority of users who did not initially manage to use the system to their satisfaction. Certainly 13 (68%) users strongly agreed that Using VOEU will enable them *to do their job effectively, and 'using VOEU it is straightforward to get to the information they needed'* with only 5 (26%) disagreeing, and one (5%) not expressing an opinion. 16 (84%) users strongly agreed that Using VOEU allowed them *to accomplish the task more quickly*, with three disagreeing. One (5%) of these users suggested that '*Need more time and practice to be familiar with the system*', and another that the '*Time to complete tasks needs to be evaluated over a longer period of time*'.

The analysis of the user's sense of effectiveness of the VOU3 (VOEU) system findings suggest that less than 16% of members of the population consider that using VOEU would NOT be of use to them in my job. These are still the minority, but the vast majority (84%) considered that using VOEU will *enable me do my job effectively*. The majority believe *that the technologies allow them to accomplish the task more quickly*.

8.9.6 Constructive comments (free text) from the users

An additional comment was about how difficult the user felt the software was to become familiar with: '*Need more time and practice to be familiar with the system'*. One user commented about how they felt *the evaluation process should be more protracted: 'Time to complete tasks needs to be evaluated over a longer period of time'*.

No additional comments were made about the user's feelings or emotions when using the software; or about whether the user felt in control, or about how easy or difficult the user found it to locate the information.

Other subjective comments included the fact that the system offers…'great potential. However some need for growth with respect to core orthopaedic topics' and 'I think it feels better as I get to use the system more'. Another user commented that we 'need more content in library/education before it can become useful on a daily basis'.

8.9.7 Conclusions drawn regarding the usability of the system

The findings of the various surveys suggest the population is representative of the prospective users both with respect to their discipline and experience for the clinical educational tools and that they consider the basic infrastructure adequate (with the caveats mentioned above) to provide for the basic framework into which the more complex media, *i.e.* surgical simulation, can be integrated. This suggests web-based service implementation is feasible. Having established this infrastructure, and returning to the concepts of parallel path management for validation and verification, it is now the task of the simulator itself to be evaluated. This will then allow the *two systems to be reviewed as a whole;* concentrating upon skills management as measured by established scoring systems.

8.10 Results of the VOU3 simulation – simulation validation

As outlined in Chapter 5, this is structured around the AIPES scoring system. The purpose of the trial is to achieve the following:

1. Support/refute the hypothesis that *'A virtual university infrastructure offers the potential of providing an effective distributed learning environment for orthopaedic surgeons'*.
2. Provide indications of which questions are useful and which are not, so as to be able to streamline the AIPES questionnaire
3. Provide an indication as to the user's preferred path for future development.

This has been achieved by breaking down the questions into three parts:

- ❑ The **AIPES questionnaire**, which is addressed in 5 parts separately
- ❑ The **Neurolinguistic programming analysis** of user perception as to the importance of different modalities of simulation (to compare the current method of simulation with the earlier version adopted in the 1st prototype simulation study).

- The **perceived difficulty** of the tasks using the NASA TLS scale (47;136). This compares VOU3 and the 1st prototype simulation as users are being exposed to the same tasks but via a different education system.

8.10.1 User's impression of the VOU3 simulation

The analysis of the simulation characteristics according to the AIPES scoring of the autonomy of the VOU3 (simulation) system findings suggests that the population represents the prospective users for the clinical educational tools. This is detailed in Appendix Tables 8.11 – 8.15 suggesting the population represent the prospective users for the clinical educational tools since they are representative in terms of experience and training.

8.10.2 NASA TLS

Adopting the NASA *TLS* scale for the evaluation of the simulation experience requires the user to rate the value of the simulation at a more pragmatic level of communication. These reflect mental, physical and temporal demand, the performance of the user and the effort expended to achieve the goal. The results are displayed in Appendix Table 8.16. The questions that are most discriminatory for each aspect of the review will be kept for future trials. Neurolinguistic analysis is detailed in the Appendix 7.1.

8.10.3 Analysis of the VOU3 simulation – shoulder arthroscopy confidence

9 out of the 22 candidates tested (41%) expressed an interest in the subspecialty of shoulder surgery. If such tools are demonstrated to provide a reliable method of training certain sub-skills such as pattern recognition, the significance of this bias toward practising shoulder surgery procedures will rise as access to simulated training environments improves, and more surgeons use them. Simulated environments may actually boost the confidence of the surgeon to perform the procedures. The significance of the system in training will not necessarily equate to the level of surgical competence that is being achieved.

8.10.4 Results of the VOU3 simulation – user simulation perceptions (C)

The VOU3 beta-tested the virtual shoulder arthroscopy environment. In Chapter 3, it was emphasized that the intrinsic advantage of designing, building and emulating minimal access surgery is that the user is in effect already immersed within the appropriate operating environment, since they normally obtain a view of this indirectly via a monitor. The display system (interface) replicates the normal colour high definition (HDTV) monitor with which all the operators are familiar, avoiding cognitive dissonance. Following 2–3 minutes of 'hands on' experience with the simulator to familiarise themselves with the system, which was using a 'genuine' arthroscope and thus had a familiar 'look and feel' to the conventional surgical equipment, users were subjected to an exercise which involved assessing the simulation itself, as outlined in (Appendix Table 8.17).

These perceptions are dependent upon the time and context of the study. With a rapidly evolving world of media that exposes the users to expectations that are much higher for image quality and convenient interfaces, it is important for users to appreciate that this is a research trial and not a system prepared for commercial exploitation.

8.11 User simulation perceptions of the VOU3 simulation

The users were asked to negotiate the route through the shoulder joint that would represent the standard navigation conducted as part of a diagnostic shoulder arthroscopy. The order (sequence of steps) of this was based upon the standard arthroscopy as outlined by Mr Bunker (Consultant Shoulder Surgeon) (150).

8.11.1 D - Simulation task performance

This involved a time-limited exercise to enter the virtual environment (VE) of the shoulder joint and navigate to obtain a good view of three structures inside the joint. These were:

- Insertion of Long Head of Biceps
- Foramen of Weitbrecht
- Inferior Recess

Another surgeon confirmed identification independently. All surgeons performed the task within reasonable time (less than 60 seconds).

8.12 Discussion

This section addresses issues on two levels: the process of evaluation, and the results generated as a consequence of using the usability and simulation evaluation measures. It prepares the reader for the consequences of these indicators. Usability testing confirms user interface effectiveness. The perceived educational value will need to allow for skill fade and the ongoing simulation assessment for performance. The system meets the tasks defined for the design in basic VE autonomy, interactivity and presence. The educational environment and scenarios can be added as the system evolves, broadening the range of training tools available. This task generation fulfils the role and assessment potential for future development methodology outlined in Chapter 1.

8.12.1 Application of the AIPES scoring system to the results of the 3rd prototype simulation

The system is charged with simultaneously providing support for the path of user perception recording (usability and simulation validation) whilst the system is itself being beta-tested for the effectiveness of the interface as part of the VOEU trials. The alpha testing as part of the VOEU usability review was not an adequate sample size for statistical significance; its purpose was a feasibility analysis of the process.

8.12.2 Relating the educational scenario to the pedagogy

This relates to Questionnaire E. The following pedagogical principles need to be addressed and are scored on a 5 point ‘visual analogue scale’ (VAS):

1. Feedback was obtained during the activity (a series of steps) that any actions or thoughts were either right or wrong (151).
2. The problem-solver has the opportunity to start a single step in an activity again.
3. Feedback during a single step will not prevent the problem-solver from accessing the following steps of the activity.
4. The problem-solver cannot proceed to the higher-level activities until he/she has successfully completed the lower level activities.
5. The problem-solver must be given the information obtained from the investigation over time, allowing reasoning to occur at every step of the activity (127;151).

8.13 Summary

The ICT skills of potential users were evaluated using the technical support book distributed to users. Continuous review supported by the client/server architecture allowed users to express their perspective and managed the user administration issues. The results of the usability trials were subdivided into the main areas of navigability and ease of learning.

The ‘pre-study’ (Version 2) provided design feedback for the final feasibility study (VOU3) and its proposed methods for data collection. It emphasized the need for greater integration of end-users into the design process. VOU3 focused upon the assessment methodology, the internal structure and the content of the multimedia educational orthopaedic modules. All of this was based around an interactive course.

Results demonstrated that the iterative design and development process benefited from user involvement. By setting and attaining quantitative usability goals it will now be possible to engineer the usable tools that may be derived from this research project as part of a rollout plan for technology implementation. The ease of testing using VOEU technologies recommends the system implicitly.

Testing versions with users early and continuously supports design iterations. By identifying users who are able to develop the system, running trials and analysing tasks, usability specifications were set for the developing prototypes. A Usability Test Report can be generated managed by the DRJ. In future significant design improvements may be measured by comparing datasets. Trainers will be able to distil the questionnaires to only use the questions that are truly discriminatory and so minimise the effort required for feedback, decreasing the incidence of user fatigue with the analysis method.

The difficult part is the selecting and organising of tasks to test and creating task scenarios that are realistic. The Zeltzer cube analysis partly demonstrated the future path of development and the proposed revision of this (AIPES scoring) for the VOU3 simulation has demonstrated a workable alternative. This also depends upon preparing the test environment. From the pilot study it was clear that a usability engineering approach and an iterative design approach were needed. By developing user profiles it is possible to select and organise tasks to test and create task scenarios (cases) that are suitable to the training grade of the trainee. The question is now how to progress from *‘results’* to *‘use’*.

9 Conclusion

9.1 Overview of work so far

The thesis has led to the construction of working prototypes of digital technologies offering an environment for familiarisation with new surgical procedures and the management of clinical case audit. This supports the orthopaedic knowledge management process. The aim is to provide a novel route for access to educational material that more closely resembles the working practice of the arthroscopist. This will support life-long learning in higher surgical training. The proof of concept was developed for a shoulder arthroscopy simulation model as an interface for the surgical trainee to access multimedia-based educational orthopaedic modules. This demonstrates a human-computer interface that more closely resembles the process of factual knowledge association during clinical procedures, moving toward the ultimate goal of seamless integration of knowledge repositories with clinical intervention, augmenting thought and action. It provides the user with context for knowledge.

The methods adopted enhance access and control of these virtual workspaces via an *eXtensible Mark-up Language (XML)*-based interface. This links the educational environment of the *Web-Based Training (WBT)* scheme evaluated in *the 3rd design phase simulation*, and the clinical data collection from ongoing trials in Orthopaedics. The underlying Virtual Orthopaedic European University (VOEU) infrastructure is designed with a broader base of data management in mind, encompassing the normal working environment of engineers and clinicians using computer-assisted surgery tools adopting dedicated interactive media.

Adaptive hypermedia associates multiple hyperlinks to the simulation interface via frames in the video sequence of the video-based surgical simulation environment. This associates the surgical experience of the trainee with the relevant course material and was positively received in the *3rd design phase simulation*. The material is prepared from actual patient operative video information, integrating with a structured surgical course model. This supports the multimedia educational orthopaedic module, generated specifically for the learning of shoulder surgery.

Resulting in an XML-based interface enabling users to communicate using material mediated for their specific needs and allowing adaptive media based upon user experience and knowledge base. This combines declarative (factual) content with feedback from a clinical (procedural) case-based training environment. By using the XML interface, it is possible to cater for the different user hardware and software resources. This includes the media content, and language of presentation, incorporated via the virtual environment simulation.

Thus trainers can construct a framework based upon already established standards, and applicability to other surgical disciplines is anticipated. This is a way of building patient-specific arthroscopy datasets *i.e*. libraries (atlases) of pathologies. This will enable the development of networking computer architectures to assist the assimilation of multiple sources and media. The breadth of data formats (structured as XML schemas) managed by the interface offers the opportunity to review data from computer-assisted/image-guided orthopaedic surgical systems embedded in an educational environment that is quantitative rather than descriptive. This will ultimately help the development of data repositories for mining, providing feedback upon clinical case management. Future work should focus upon the process of patient data collection and refinement of the image database production. Before consolidating these conclusions, this chapter explores the impact of this work upon surgical simulation.

9.2 Introduction

This chapter starts with a discussion of the results of introducing a web-based (flexible and distance learning) training environment to surgery. The clinical trial of the effectiveness of a web-based training infrastructure provided adequate information for surgical authorities to explore this new medium as a major vehicle in education. By outlining the proposed methods for the introduction of the web-based training system in conjunction with client-server architecture of the simulator system, it becomes possible to incorporate the new technologies within our major central institutional educational infrastructures without causing significant disruption to current practice: evolution rather than revolution. This is likely to be provided through major training centres with good networking connections. A *'store and forward'* approach is likely to be adopted. By comparing the thesis statement with the clinical results, it is clear that the system is unlikely to be viable in its current form, but as a demonstration prototype it provided a tool to evaluate the principles upon which the next generation of surgical simulators is being designed. The results highlighted the need for cultural change within the field of surgery in line with the development of tools to meet future educational demands. This led to the development of the Orthopaedic Computing course, termed *'ICT Skills for healthcare professionals'* and the preparation of the Basic Computing for surgeons course.

With an integrated patient-centred information management system becoming the backbone of hospital information systems, *web aware* staff are more likely to use such systems as they become more familiar with the networked infrastructure in hospitals. This will lead to integrated patient management systems supporting simulation, since ***surgical simulation cannot be considered in isolation, but must be seen as part of the integrated record / performance measuring system*** that may direct the evolution of surgery in this century. The evolution of the medical networking infrastructure will dictate the progress of this. The need for more sophisticated database construction is explored.

9.3 Discussion of the results of introducing a web-based (flexible and distance learning) training environment to surgery

The VOU3 trial of the effectiveness of web-based training infrastructure provided adequate information for surgical authorities to explore this new medium as a major vehicle in education. By introducing web-based training (WBT) in conjunction with client server architecture of the simulator system, it becomes possible to incorporate the new technologies within our major central institutional educational infrastructure. This is likely to be provided through major training centres with good networking connections such as District General Hospitals. With the introduction of home accessible broadband (0.5-1 MBps) connections there is a reduction of geographical dependency upon hospital sites. A *'store and forward'* approach is likely to be adopted when considering more pervasive access for data management and collaboration. Since access to professional educational materials is tax-deductible, it is not unreasonable to consider broadband facilities as a major expense.

9.3.1 Consideration of the thesis statement in view of the clinical results

VOEU has to be considered both for its ability to integrate with current work practice and hospital information systems, also for the potential of simulation in learning.

Hospital information systems: The integrated patient-centred information management system is becoming the backbone of hospital information systems along with teleradiology imaging systems (152). This could set unrealistic expectations for those who would compare corporate database provision with research tools still in their infancy. In the long term, though, these systems will play a key role in underpinning the development of a networked infrastructure in hospitals. The educational, cultural and social integration is ready for simulation, though as yet, the simulation is not versatile enough for performance measurement: it is solely for research purposes, though this may evolve through time.

A set of good practice guidelines was prepared at the request of the NHS Executive in consultation with the Joint Computing Group of the General Practitioners' Committee of the British Medical Association and the Royal College of General Practitioners. Its primary purpose is to provide a professional framework for the legitimisation of electronic patient records in general practice. It is also intended as a source of authoritative advice for those general practitioners who keep computerised patient records or who intend to transfer their record systems to computer.

The reality is that with the current systems there is still a *human in the loop* and, although they may be set up to filter and compile information independently, they remain under human supervision and control. VOEU is primarily designed to document, but with messaging to support this process. The inputs can be refined by the VOEU human-computer interface. The outputs are interpreted by users better versed in the technologies by virtue of the ICT skills book, but the pivotal processes remain the navigation of the educational spaces that are potentially enhanced by using embedded simulation. The intention is to deliver the appropriate message (usually as a document) to the recipient as effectively as possible.

The impact of simulation on surgical learning: The embedded simulation has been evaluated as a useful usable learning tool. This is not necessarily proof of its effectiveness in learning. To prove this would necessitate a blinded controlled trial. This is beyond the scope of this thesis, but it is the intention of the author once a broad enough range of simulation worlds have been generated. Different pathologies can then be demonstrated and the knowledge tested.

This will not only test structured knowledge, comprehension of tasks and application of techniques but also, according to Bloom's taxonomy (31), it should aim to demonstrate ill-structured learning goals, such as analysis of (virtual) environments, synthesis of a clinical management strategy in response to findings, and evaluation of results of interaction. It is also potentially possible to reverse the role of the interface by developing the simulation as an interface to navigate through the content. Vital to the potential success of this is the increasing role that the customised 'off the shelf' systems (COTS) play. Any future evaluations are likely to exploit technologies primarily developed for the 'gaming' markets but with more sophisticated analysis of the results.

The most important demands made by simulation will be on content generators producing the video sequences with embedded links and dedicated authoring tools. Adequate training will be essential for the implementation of this, and also suitable tools for the media preparation.

9.4 VOEU project impact upon surgical simulation integration

Fundamental to the success of the simulation model is the integrating of learning and visualization technologies in Orthopaedics. This will be achieved, in part, by establishing 'web service' derivatives of VOEU, based upon the tools in Chapter 7.

9.4.1 The end users

In Chapter 2 it was argued that surgeons are mobile well-educated individuals whose work demands excellent availability of educational material that is up to date and focussed on their particular learning situation. This reflects the need for life-long learning (continuing professional development) and refresher material as well as the *'just in case'* archives. New material is being collected constantly and this needs to be properly evaluated and integrated into the learning infrastructure appropriately.

The developments in VOEU provide the process of data collection and retrieval. The learning environment and simulation tool are therefore symbiotic. The VOEU system has been developed in a way that will support more sophisticated media including virtual environments.

The XML (108) foundation presents the user with data in a form that is interpretable for their level of knowledge and experience based upon their personal profile, and the system is fully scalable with respect to the number of collectable cases. It can be used to record a permanent archive of clinical procedures. Because of the sensitive nature of the material, the security issues are addressed at the server level and consent must be obtained from the patients for their case material to be included, as was the case in the trials for the *1st design phase* simulation and the *3rd design phase* simulation.

Part of the educational strategy relies on the evolving *'virtual university'* concept to allow for the updating and upgrading of educational material in light of new results, using analysis of incoming data from ongoing clinical trials for the evidence base. This acknowledges that a virtual university is a living infrastructure that evolves with time, due to changes both in its underlying philosophy and staff.

The construction of metadata standards for the core components that abide by already approved standards, such as the Dublin core (153), *Learning Object Metadata (LOM)* and *Information Management System (IMS)* (154), leads to a philosophy where components are 'living entities' whose survival within the university infrastructure relate to their relevance to users. Applicability (the ability to apply the technologies for other applications) and expandability (68;111) are vital characteristics of these components.

The conclusion drawn in this chapter is that VOU3 demonstrated successful integration of components rather than using proprietary embedding software as suggested in Chapter 2. This means that the boundaries between the educational media are economic and social ones rather than technical.

9.5 Applying the learning model to training in a virtual university

The vocational training (155) of anyone whose work involves clinically invasive procedures requires the best from that individual with regards to decision-making and psychomotor skills. The development and monitoring of these are therefore vital. Ultimately, simulator-based skills training will become integrated within the broader framework of clinical governance (the quality assurance of clinical service delivery).

The methodology of previous task analysis for evaluation of surgical skills training was based upon the practical limits of what could be achieved with the available equipment at one site and the expert analysis of tasks (61;139). The results were inconclusive with a small number of trainees, as would be anticipated using descriptive statistics. It is therefore clear that a larger number of trial candidates are required. This is being addressed by the development of the virtual observatory linking to the DRJ within the Virtual Orthopaedic European University (119).

9.5.1 How users benefit from customisation in a virtual university

The appropriate knowledge category depends upon the clinical role that the individual plays and their stage of training. It is the responsibility of course convenors to ascribe various XML tags to learning objects to match the syllabus for the grade of the trainee. This allows for focused presentation of material to the user group. This is achieved by providing 'look-up' tables for each grade of trainee. Improved functionality results in increased complexity of the interface by using a derivative of Structured Generalized Mark-up Language (SGML). The XML abides by standards being evolved for this domain by IEEE and, where possible, the international organization, Health Level Seven (www.HL7.org). This will ensure greater compatibility between the interfaces that are used for the control of surgical procedural trainers (simulators).

There are potential pitfalls, though. The most significant of these is security. The system is stable and likely to display good survivability characteristics (140), though the open nature of the network and its potential application for providing communication portals between the machines can provide a vehicle for corrupting software (viruses). For this reason the ftp functions previously used in the Java control interface of the 1st prototype simulation and the file management for the learning objects now have to be handled differently, with *SQL* for the former, using *.NET* and email or dedicated uploading procedures for the latter.

Personalized libraries can be built through VOEU Integration driven by the assertion that gaining and retaining knowledge are personalized processes that need to be individually customisable. The first stage is to match the user to a specific group who require core knowledge for their work. The assertion that different 'grades' of staff have differing job descriptions requiring demonstration of differing levels of knowledge is generally accepted. Bearing in mind the perspective of the users, this involves educationalists, engineers, and clinicians of differing experience.

Users are likely to develop their own preferences for how the system is used. Data are therefore collected and stored centrally to allow the user to modify their preferences according to the level of their expertise, matching the presentation of the data to the user prior knowledge and experience as defined by the user profile.

9.5.2 Role of the dynamic review journal in beta-testing

The DRJ supports continued *beta-testing* to refine the applications so that they will be available to a wider audience. This can be adapted by adding new virtual environments and building a repository of cases that can be used for paradigm (problem-based) teaching. It will allow central collection of adequate data for both interim and long-term analysis of surgical progress. This will be formally evaluated in Version 4.

The framework will be enhanced by the development of patient-specific databases of information, including the potential to build patient specific virtual environments for practising surgical procedures. This leads to the opportunity to create a range of different safe virtual testing environments for the design and evaluation of new surgical tools. The URL for this is: http://voeu.ecs.soton.ac.uk/library/doc.aspx?docid=?

The Dynamic Review Journal (DRJ) is part of the permanent review process. By establishing the web-based archive, technical and medical material will be reviewed and updated. Future teams may then be able to continue this work, building upon this project. The aims of the DRJ-framework in beta-testing are:

1. To enable contributors, such as scientists or clinicians, to prepare publications for their specific results of research.
2. To enable the development of a current (dynamic) medical and technical library in the domain of Computer-Assisted Orthopaedic Surgery built in units that can be combined and decomposed in meaningful ways.
3. To provide researchers with a publication model that supports collecting and sharing comparable data and presenting it in the form of standardised learning objects.
4. To define a standard that is simple yet extensible to all areas of surgery and medical scientific reporting in general so as to be most easily and broadly adopted and applied in the current medical literature environment.

9.5.3 Profile construction

The main inputs are going to be from VOEU's component systems that provide data on an individual surgeon's performance. These will include training simulators and the intra-operative tracking data. Producing the necessary output for individuals and their tutors to appreciate the real strengths and weaknesses of performance will require the DRJ to provide a range of data processing visualisation tools for the end-users to be able properly to appreciate the data within the results obtained.

9.5.4 Data output presentation

Standard measures of performance have been adopted already within the UK. These will be incorporated along with new measures that are yet to be tested. Since the work is research and does not impact upon the present performance standards of trainees, by refining new standards using the data collected unbiased reliability can be demonstrated prior to actual use. For trainees who already have an established clinical practice, the system should allow for *'near real-time'* auditing of the cases carried out, providing the most complete and accurate record of outcomes as measured by the recognised systems for recording this, including both general, such as *SF 12* and *SF 36* and specific, such as the *Oxford Hip Score.*

Users need to be able to customise the output into a format that they can easily recognise and relate back to their everyday clinical practice. This may lead to the most recent results becoming available in the everyday clinical arena, such as the presentation of accurate outcome data to patients, when obtaining their consent for surgery, as to the relative merits and risks that they are about to take.

Patients may find this useful. Surgeons may be somewhat concerned by such an informed population. There are occasions when such information is of limited value due to the wide range of parameters affecting surgery that cannot always be taken into consideration. At each stage of system rollout the old adage of *'set, hold, trim'* will be adopted, setting up a new component or system, holding it with a design freeze for six months to a year to become established, and then upgrades (trimming) will occur so that the next version is able to accommodate the wishes and demands of the users. This is an iterative implementation process shadowing the iterative design process. This has implications for journal design.

Peer review: Traditionally scientific and medical journals have been respected for the quality of their material that has undergone critical independent peer review. The DRJ must align with this process if it is to gain equal credibility. Surgical Review Boards might ensure that a suitable peer review process is managed.

Open access and 'popularity' selection processes: There is the potential for the introduction of an open democratic selection process using this system. Any individual can include results in the DRJ, but only other users, including course tutors, would regularly recommend those articles that are considered relevant. This provides a selection pressure upon material. It ensures that all contributors get a fair opportunity to publish.

Such a system is only possible for experimental data that has been peer-reviewed, where the methodology and integrity of the scientific process has already been validated. This is truly 'Evidence-Based'. The potential flaw is that 'interim results' may introduce bias that may mislead. Users may not be fully aware of the process involved. It is a system that reliably monitors the surgical performance when tracking systems are employed, but may not reliably measure the all-important outcome measures.

Presentation medium: Users may wish to adapt the templates to meet their specific demands. Where possible, standardisation is encouraged, since it will assist the searching systems. It is however recognised that the basic templates will not be suited to all applications. The aim is to provide a system where the user is able to develop the infrastructure to meet the specific needs of his/her field, so individuals must be able to modify templates.

Principles of confidentiality – relating to the DRJ: With the need for patient-specific data in educational content, users need protocols for peer review and anonymisation of demographic information. Each country has its own mechanisms to assure protection of data stored regarding the rights of individuals. The data protection act for each country must be adhered to. Transferring data from a clinical investigation (*e.g.* MRI scanning) for the preparation of training material requires a process of anonymisation. It is necessary to ensure that this data is not normally traceable to source, except by a recorded responsible party such as the primary care doctor, since new information may have an impact upon clinical management. The principles of patient confidentiality are outlined in Appendix 9.1.

9.6 Definitive messages from VOEU

1. Virtual University infrastructures are essential. From tiny acorns, great oak trees grow, and so too with the university (XML) infrastructure. These tools should be free at the point of delivery for users who are likely to be granted access via institutions and societies.
2. To qualify the data, VOEU established schemas that may ultimately be integrated into an educational schema. VOEU's role as a knowledge construction tool, integrating concept mapping with supporting data repositories, offers a powerful base to support research, clinical and educational tasks.
3. XML plays a significant role in system integration. This allows different simulation systems, each with the capabilities to provide training and evaluation in

specific surgical sub-skills, contributing data for evaluation. This provides different grades of user with the appropriate tailored information that they require.

4. The database will host the relevant learning cases collected clinically that will then be used as test cases for validation purposes of surgical simulators, providing automated support for the validation and verification phase. This addresses the issue that: "*One cannot properly assess the performance without first knowing the performance of the population who are testing it*" by setting the perceived performance of the system and then comparing changes to this. Simulated virtual environments can then be designed into the interface as portals.

5. The evaluation of simulator performance should be progressively refined to ultimately provide a well-founded system of assessment for a large population of surgeons. Distributed computing architecture will be required to develop this important database.

6. The aim of future work is to ensure that these devices assure compatibility between the intra-operative tools developed for the training and microsurgical operating environments and also the extension of the archive of cases for testing of new tools. The simulations will be linked via the adaptive hypermedia system to act as portals spiced with the particular links requested by the user, integrating the visualization process with the learning content.

7. The educational environment of the Web-Based Training (WBT) scheme and the clinical data collection from ongoing trials in Orthopaedics are linked. The material is prepared from actual patient operative video information, integrating the structured surgical course model with the multimedia educational orthopaedic modules, generated for the learning of shoulder surgery.

8. A similar '*Look and Feel*' fits with a cohesive integrating collective policy that extends beyond the hospital and the region, indeed the nation, to the heart of the healthcare infrastructure - the person. It emphasises the principle that to make human computer interaction better we are adapting both the human and the computer. The system displays admirable qualities such as non-repudiation of content and presentation of material in a context-dependent setting. This properly exploits the potential of adaptive hypermedia (AH), with a customisable front page for the individual. By cross-referencing with embedding of tools such as commercial reference-managing software

e.g.: Reference Manager™, PubCite™ and *EndNote™* this could become a very powerful knowledge management resource.

9. Future development of context-dependent searching will adapt the keywords lists to meet the specific demands of the user with regards to their personal portfolio and educational needs. By introducing the concept of Critically Appraised Topics (CATs) into educational modules this integrates evidence-based medicine into the learning material. By searching a link to personal reference link bases, the ease of keeping the knowledge base current and appropriate has been confirmed. The underlying technologies linking this are XML and SQL-based. These technologies should be adequately robust to be managed with split-site server farms. A web service infrastructure would better provide this.

10. In addressing the issue of integration it is necessary to provide an answer to the question '*What is a digital library?'* for any components will ultimately be part of one. The University of California refers to a digital library as 'systems that support the collections of the University' (156). The users should therefore broaden this term from the media collections themselves, to include other information *e.g*. possible experimentation, analysis, and collaboration, where both information dissemination and formal publishing are equally important. The DRJ has been integrated and used, providing an example of a system which erodes the barriers between the areas of publishing, dissemination, discussion and education.

9.7 Future work in virtual university technology implementation

The strengths and weaknesses of alternative simulation systems that are currently available, compared to the ideal parameters for a surgical simulation medium, are discussed below. The concept of shared systems in complex operating theatre environments leads to the need for integrating many hardware and software components and thus has major implications for interface design, emphasising the need for standardisation of systems. Though this addresses the simulator's environment, the important question of developing the concept of video and graphical integration remains and has been taken forward by others *e.g.* University College London. Before the exploitation of the real potential of such video-generated virtual environments, it will be necessary to integrate techniques of image formation from both video and graphics (157;158).

By establishing an integrated policy to incorporate computer-based learning within the existing educational infrastructure, the need for an integrated approach between regulatory authorities, research institutes and industry has been clarified. This relates to the advances in new surgical tool technologies. *Policy Formation Issues* arise, reflecting the social, economic, and political pressures that are brought to bear. Thus, future plans regarding networking and regulations emphasise the need for interaction with the surgical authorities. The strategy for uptake by authorities involves *The Royal College of Surgeons of England* and the *British Orthopaedic Association.* This is explained by the need to achieve this political and social integration with conventional surgical educational infrastructures.

To enhance healthcare there should be direct or indirect benefits to patient care. The predicted direct benefit of these systems is in the form of improved surgical training and thus a reduction of direct risk of morbidity as a result of clinical error. Indirect benefits result from a more cost-effective training system.

9.8 Surgical users' perspective on implementation

The surgeons (or any professional for that matter) need to a degree to display the following characteristics. They need to manage information, *i.e.* have the knowledge to hand or be able to gain ready access to it. They need to be capable of decision-making and planning based upon whatever information is available at the time. This relates to the tools and equipment that they also use. They need to have the necessary skills for communication and to use the equipment, and most difficult to measure, they need the appropriate aptitude. This includes flexibility, confidence, persistence, patience and courtesy. Other instruments for assessing these qualities exist.

'Virtual Reality' technology cannot currently solve all these training issues but it should allow surgeons to rehearse the skills of tissue manipulation, excision and repair, without putting the public at risk. Although the set-up cost in terms of development may well be high, in the long term, using client-server and web service tools, it may be more cost-effective than any other method. Minimal Access Surgery (arthroscopic, laparoscopic, thoracoscopic) in particular lends itself to computer simulation, for the surgeon is working from a monitor that can easily be replaced by the computer screen. Most conventional arthroscopy (single camera) renders the surgical field two-dimensional, allowing the surgeon to be immersed in a 'desktop' simulation, whereas simulation of open surgical techniques will require stereoscopic head-mounted displays and kinaesthetic gloves for a fully immersive VE.

Getting users up to speed (extra learning vs. time available) will depend upon trainers and trainees agreeing realistic timetables, supported by Individual Learning Agreements (ILAs). From then on, maintaining awareness of technologies will be managed through the continued updating of users and increasing awareness of technical solutions available to them. Building the system by personal or team contribution is by far the best way to gain support for the new technologies as users extend their role as stakeholders, taking pride in the resources that they prepare for others.

9.9 Future simulation development methodology

The preparation of a multidimensional video database is a novel approach to surgical training, and in effect represents a progression from conventional 2D film to 3D. This allows for improved depth perception, highlighted in the task analysis, to include 4D, that allows for the modelling of both orientation and navigation skills that are essential for the trainee of minimal-access surgery to acquire. By working with video, one has the advantages and disadvantages of adopting what is essentially a *'completion backward principle'*.

This means that the recording of a video image provides all of the necessary visual qualities to produce an accurate representation of the environment. Use of video technology alone is limited. The density of the database is inversely related to the greatest distance between adjacent frames. There is no fixed distance between the *Virtual Observer* and the *Virtual Object*, therefore calculations relating to the cosine of the arc as a way of measuring the relative distance between frames in rotation are impractical. Other combinations of recording parameters for granularity were considered with the indicated size of database required to record these. More detailed analysis of this problem is published elsewhere as a research report (138).

Clearly the use of video is not a panacea for all simulation, since it will not give useful information concerning all the qualities (parameters) of a virtual environment. It can however be used it for the acquisition of visual data qualities/values that, if accurately registered, may be incorporated into a graphically generated simulation environment (36;64;87).

For this to occur the system depends upon the accuracy of registration. By virtue of the current generation of surgical instruments it incorporates a need for 4D recording, due to the angle of the lens of the arthroscope recording an image positioned on an oblique axis of 30^o so that for any ***x, y, z,*** co-ordinate a full 360^o range of α co-ordinates (representing the rotational component) potentially exists.

The recording of all of these potential positions necessitates a quartic increase in the number of frames that need to be recorded. It becomes clear that this system would very soon produce an unmanageable number of frames, so that they would not be able to be recorded manually. It does however generate a high quality template for photo rendering applications. In particular the use of non-linear morphing of images may help to overcome this issue (158).

9.9.1 Developing the concept of video and graphical Integration

Before the real potential of a video-generated virtual environment can be exploited, it will be necessary to integrate techniques of image formation from both video and graphics. This could be performed using hardware and software, but in order to maintain the "*Willing suspense of disbelief*" it must be done in real time, *i.e.,* 0.05 seconds per frame. One of the fundamental problems with simulation of a surgical environment is that one must consider the temporal dimension as being linear, fixed and unidirectional. *i.e.* time does not run backwards for the purposes of this work, and must be represented within +/- 10% of the actual time reference, in order to be convincing to the observer.

Consequently, it is fundamental that any environmental change is immutable in that, should the scene change, the future frames must accurately represent that change. This is the Achilles heel of a video-based system.

To be of benefit in surgical training the simulator must provide visually accurate representations to provide material for pattern recognition of pathology. It must provide spatial representation of the possible dimensions of movement to provide interactivity and seamless integration of objects within the field of view.

The potential problems of video / graphical integration lie in the potential inability to achieve interaction between the two formats. This might be better considered as a way of rendering a graphical world in photo-accurate rather than photo-realistic images. The normal issues of registration and object collision avoidance arise. Because video frames are photo-accurate, the image already has the salient properties 'calculated' and represented. Analysis in Table 9.1 presents the strengths (and weaknesses) of the respective systems:

Aspect of Model	Simulation Features	Video Based Worlds	Graphics Based Worlds
Observer	Position of the virtual observer	√	√
	The perspective projection	√	
Lighting	Colour of objects	√	
	3D modelling*	√	√
	Illumination	√	
	Reflection	√	
	Radiosity	√	
Geometry	Hidden surface removal	√	
	Geometric representation of 3d curves and boundaries	√	√
	Transformation to a limited degree.	√	
Physical	Dynamics		√
	Animation		√
	Free-form deformation		√
	Detection Elastic collisions		√
	Crude collision		√
	Modelling transformations		√

Table 9-1 Development of surgical simulation requirements

The potential solution lies in the method of extracting key photo-rendering information from salient stills and applying appropriate algorithms (158). One method would be to use the labelling software (Figure 9.1) that the author previously developed with Dr C. Setchell in University of Bristol, so that the virtual environment is 'auto-rendered', including adding links to a linkbase as well as photo-rendered.

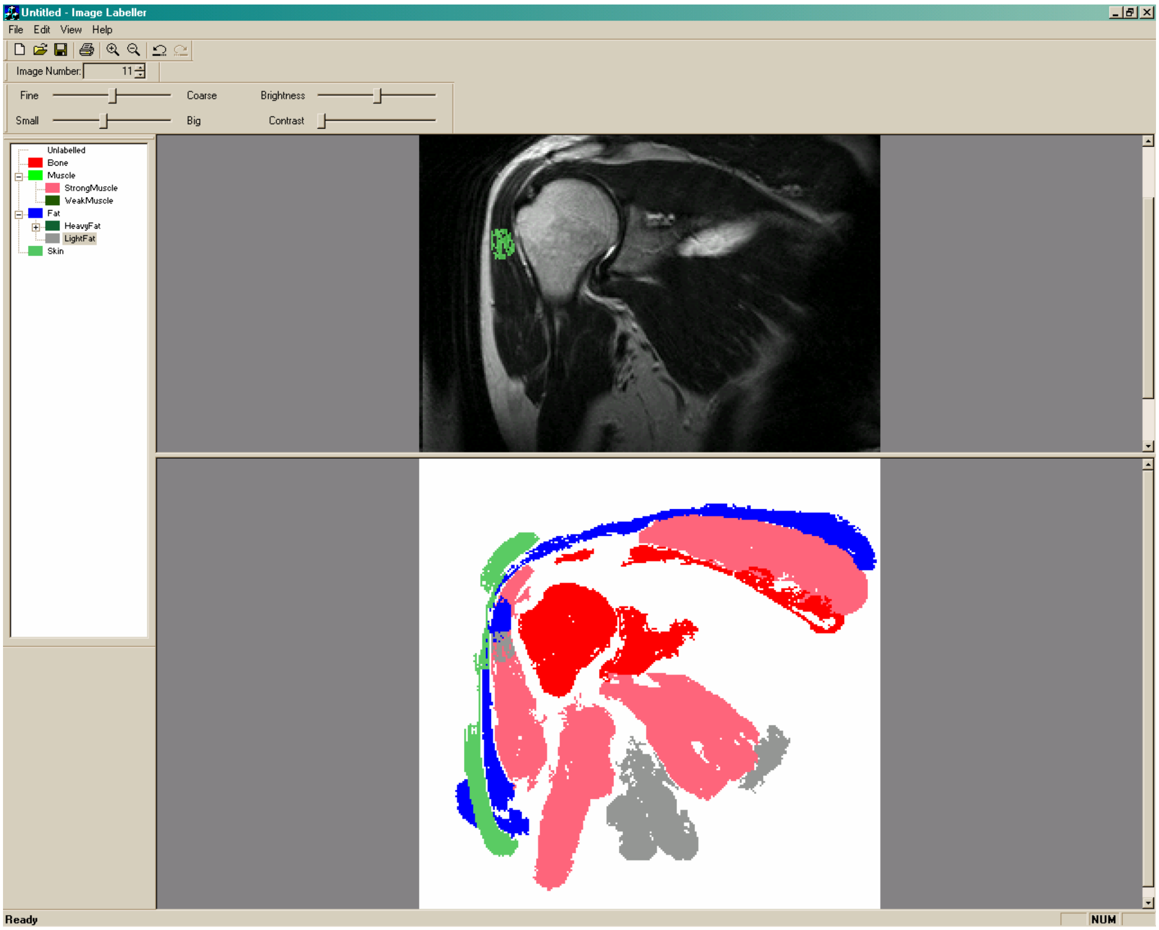


Figure 9-1 Labelling tissues for rendering using video/ haptic or hyperlinking

9.9.2 The Client/Server format

Decoupling the processing into a client / server format running over an ATM network improved the system performance ten times in comparison with the prototype system. This was largely due to particular properties of the video codec so that the serving computer was no longer processing graphically but merely storing and retrieving (video) data.

A single server could manage ten training seats on the network, which may be local or remote. A typical installation would deploy a broadband network with ‘guarantee of service’, e.g. ATM or fast Ethernet (when *IPv6* is implemented) locally, run one or two servers and provide a laboratory of training seats located in the same physical area, distributed throughout the organisation, or to remotely operated sites.

The advantage of having a centralised single *virtual surgery* server is immense. The economics of scale apply. Once the system is installed it can be used by other disciplines within medicine as applications are developed, *e.g.* Teleradiology using a

PACS system (159). Ultimately, as with other aspects of healthcare provision, it is likely to represent a balance between public and private partnership that central funding (*Department of Health* funding) would be key to the success of a reliable training system providing fair access to the surgical trainee population.

9.9.3 Transfer to web services

Web services offer limited skills training potential, employing tools such as embedded Duplex video linking. They do however gain the advantages of true web services, supporting a ubiquitous computing philosophy with central systems administration including security and content management. The author argues that it is only by achieving such advantages that the uptake will come through empowerment of the end-users. Bandwidth is important. This is likely to be around 500 kBs^{-1} initially progressing to 1 MBs^{-1} as users will be bound by personal cost considerations.

9.9.4 Specific constraints of VE evaluation

The profession should encourage the development and introduction of surgical simulation. Implementation must accommodate the existing infrastructures for currently these are basic and higher surgical training. These are based around regional courses and it is necessary to ensure that such regional needs are met. At present there are no regulations for the authorisation or use of surgical simulators, nor standards set to aid the introduction of them nor digital educational support.

Demonstration of the effective usability of VOEU means that it is ready for the implementation of phased development using the networking infrastructure (SuperJANet III). Browser, workstation and desktop performance are improving in accordance with Moore's Law, and flexibility in building distributed architectures allows for a wider range of options. Simultaneously it is necessary to consider both the limitations of VE models and the sub-skills being taught. Will the systems fulfil their primary role of teaching trainees correctly? Ultimately the ability to obtain suitable test populations for studies and arrive at a cost-benefit analysis in order to gain political support for the applications is vital.

9.10 Medical educationalist's perspective upon implementation

9.10.1 Life-long learning

The Department of Health's document "A First Class Service: Quality in the New NHS: 1998" is one of the foundations upon which the most recent NHS reforms are being built. It introduced the term *"Clinical Governance"* as

"The process by which each part of the NHS quality-assures its clinical decisions. Backed by a new statutory duty of quality, clinical governance will introduce a system of continuous improvement into the NHS."

There are three main areas where improvements are to be generated: the setting of standards, the delivery or maintenance of standards, and the monitoring of standards. On an individual level life-long learning is a means of maintaining personal standards for all health care workers. To quote the Department of Health,

"Life-long learning will provide NHS staff with the opportunity to continuously update their skills and knowledge to offer the most modern, effective and high quality care to patients."

An essential part of life-long learning is the ability to obtain reliable feedback from journals. With regard to attaining ICT skills, it is up to individuals to accept the responsibility and secure this training for themselves - the options being *'self learning'*, *'in-house'* or external training courses. *"life-long learning"* is the key, as with other aspects of medical education. To accommodate the clinical duties, a structure has evolved where surgeons attend either specific courses or local-level training meetings. The specific courses are 1 to 3 days in duration, as an integral part of the Continuous Medical Education (CME) / Continuous Professional Development (CPD) infrastructure.

9.10.2 Potential impact of a DRJ upon curricula

The DRJ is likely to impact upon the curriculum in the domain of the Continuing Medical Education (CME) points system. The curriculum therefore has a bearing upon the development of the DRJ, in line with the advice of Jolly *et al* (160).

"A curriculum needs to be both explicit and designed by a consensus involving wide representation from interested parties, and taking particular cognisance of the uses to which skills and knowledge gained in training will be put."

9.10.3 User requirements for continuing medical education

Each *Certificate of Completion of Surgical Training (CCST)* Board will provide standards for examinations to complete the orthopaedic fellowship. It is then the responsibility of the board also to set standards for maintaining this proficiency. Various approaches are currently adopted and detailed analysis of these is outside the domain of this project. However, users will need to participate in some kind of Continuing Medical Education scheme.

If this is to include appraisal of peer-reviewed literature, then VOEU can offer a system of assessment of this by allowing submission of answers to questions that will be set by the authors of the papers presented, which can then be marked. It is necessary for the questions to represent the knowledge that the author expects to impart to the reader and to be judged fair and relevant by the reviewer of the paper, since such questions will be set during the initial review process. Preparation of one such Multiple Choice Question (MCQ) may be deemed an essential part of the submission of the paper and so be included within the template.

Adequate review of the literature would allow the reader to collect a score which would allow them to submit the literature review as part of their total for CME points, stipulated by the Education Board.

9.10.4 Integration with a regional training infrastructure

Future plans regarding networking and regulations emphasise the need for interaction with the surgical authorities. The strategy for approaching the uptake by authorities such as The Royal College of Surgeons of England and the British Orthopaedic Association is one of dissemination through committees and organisation with a clear ICT focus, such as the Journal of Bone & Joint (British) IT committee and the British Medical Association CCST committee.

In future *supervision of training* can be managed using *'store and forward'* techniques for relevant documents tailored to and supporting the individual learning agreements. This will provide adaptability to local, regional and national demands. It will integrate clinical exposure within the curriculum ensuring that the syllabus is adequately covered.

Whilst the intention is to collect data that will provide useful educational feedback to surgeons, collecting data for evaluation of the simulation is a two-edged sword. The other edge is monitoring surgical skills performance that will eventually provide a quantitative performance marker for the surgeon. The critical factor is where these standards need to be set. It is already agreed that single measures are poor indicators of surgical performance (12;13;161;162).

This will become a political as well as ethical dilemma that will need to be met head on. Should the policies outlined in the previous sections be implemented, progress to a new era of quantitative as well as qualitative assessment of surgical performance will become accepted. When forming policy regarding surgical simulation for training one can reflect that much has been stated about the end-user practising surgical skills using a virtual environment simulation. The standards that can be achieved currently fall short of those required.

9.10.5 Joint committee for higher surgical training (JCHST)

The joint committee on higher surgical training of the United Kingdom and Ireland acts via a specialist advisory committee for each surgical speciality. This committee supervises higher surgical training assessing competencies required for a competent surgeon, but not directly related to the CCST. These include:

- Communication
- Knowledge of basic and clinical sciences
- Clinical ability (this includes history taking, clinical examination, interpretation of physical signs, ability to arrange investigations and interpret those results).
- Decision making and judgement
- Surgical skills and manual dexterity
- Post-operative management and management skills
- Teaching and learning skills
- Personal effectiveness - time management, team working, leadership and ethical approach. This includes reflective behaviour, maintaining self-criticism and development.
- Research skills and critical analysis of data

The process for continued review during training has therefore been formalised in the form of a regional in-service training agreement (RITA). This can record one of three outcomes, satisfactory progress, and recommendation for targeted training or recommendation for intensified supervision/repeated experience. This implies a conveyor belt process where trainees need to maintain a certain rate of progress, rather than a specific level of competency. It has, however, been proposed that there are really three levels of competency (163). The argument is for independent external assessment managed annually by the *Postgraduate Medical Education Training Board (PMETB)* to review individuals and decide whether they are competent to progress. At the end of four years of Higher Surgical Training, there will be an assessment of this core training leading to a certificate of completion of training (CCT), which may contribute to subspecialty training selection of the trainees.

Whichever system is finally adopted, it is likely to evolve from the current RITA system. For this reason, the form structure has been incorporated into the VOEU training system, allowing flexibility to adapt to political changes (see Figure 9.2). Both the new deal, (154) and the Calman training programmes (155) have resulted in the significant reduction of the time available for surgical training. This is apparently necessary to bring the UK system in line with European Working Time Directives (EWTD). Demonstrational competency will become more significant. Ethics and professionalism depend upon competency. The only benefit of time-based training is political expediency.

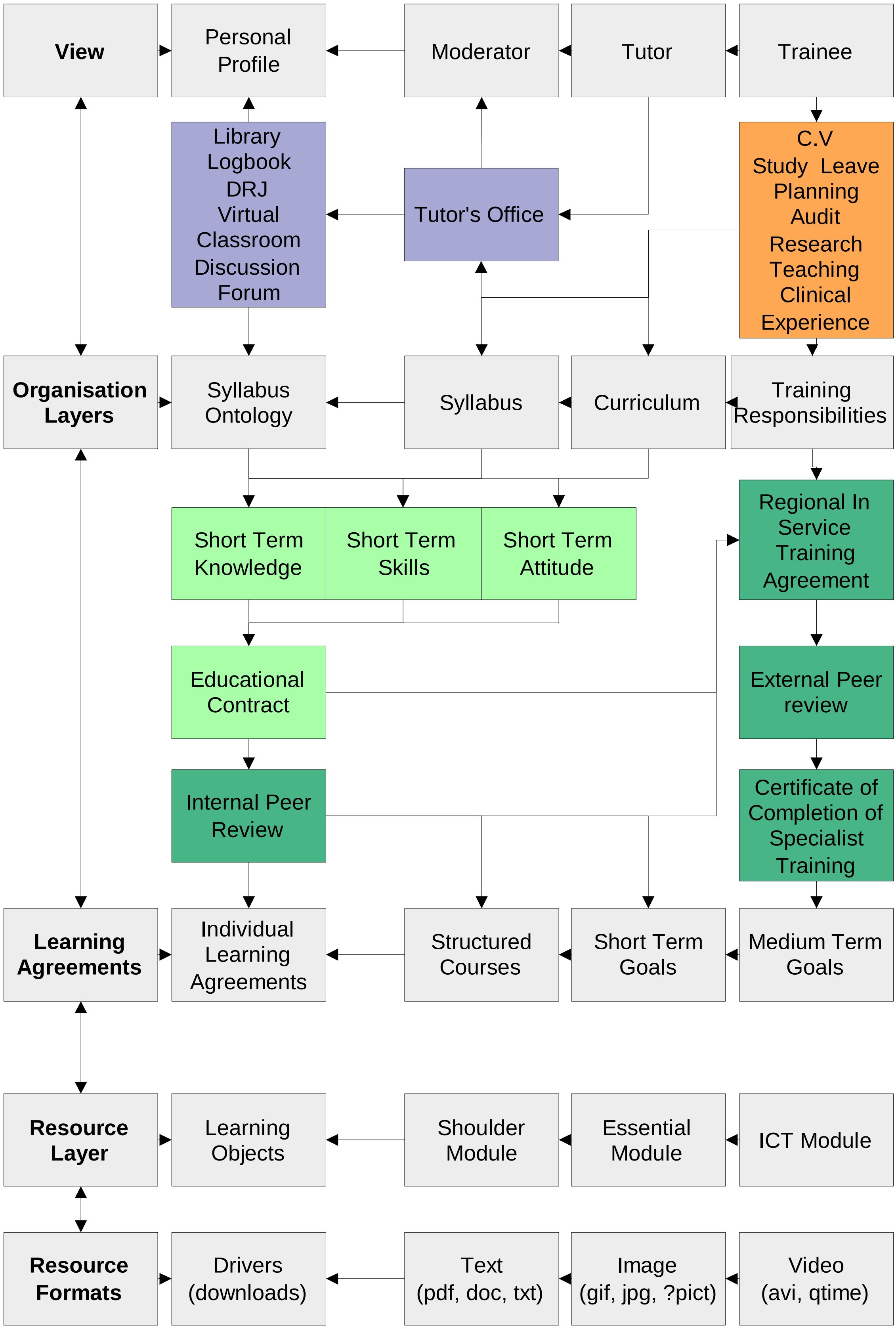


Figure 9-2 Integration of the conventional educational infrastructure using adaptive hypermedia technologies

9.10.6 *Guidelines for virtual university construction*

In healthcare simulation, guidelines should be adhered to when outlining the methods of evaluation of simulation (virtual) environments. This could be based upon the following principles:

- All systems should be upgradeable.
- All trainees should have equal opportunities for using these emerging technologies, appropriate to their training.
- Simulator application should not be limited by the platform.
- Simulators should be accessible remotely so that geographically isolated surgeons can use them either using broadband networks, or at local centres.
- Standards should be set, with increasing CME credit being awarded for ascending complexity of simulators. This can be formally integrated into the *Individual Learning Agreement.*

Surgical simulation is still in its infancy. Progress is slow because there are several small and under-funded groups tackling the same problems independently throughout the UK, Europe and the USA. Higgins and Satava (104) suggested a timeframe for the adoption of widespread use of endoscopic simulators in the USA by 1999, realistic simulation of hand manipulation of organs by 2000, simulation of two handed ligation and suturing by 2002, adoption of simulators for certification by 2005 and full integration of simulators into surgical training and certification by 2010[31]. European efforts are attempting to catch up, requiring action from the profession, the institutions and the National and European governments. VOEU is just one aspect of this.

9.11 Institutional perspective upon implementation

Preparing the populations of users: Ultimately, as user skill levels improve and designers and developers will be able to work from an 'end-user' *concept library* that supports new technologies, it will be possible to employ shared systems in *Complex Operating Theatre Environments*. This principle leads to the need for integrating many hardware and software components, and thus has major implications for interface design.

---

[31] Whilst this is 'ambitious', progress is being made with simulation being incorporated into dedicated training courses. An example being the San Diego Shoulder Arthroscopy (SDSA) course www.shoulder.com/meeting.html which is held on a regular basis in California.

Establishing digital healthcare networks: Ethical demands upon establishing a computer-based surgical simulation infrastructure need to be considered, as must the need for an integrated approach between regulatory authorities, research institutes and industry. In particular, this relates to the advances in new surgical tool technologies. UK implementation of the *National electronic Library for Health (NeLH)* (164) is moving tentatively toward this with the *'Athens'* services providing a route of access for individuals and a development policy that intends to include personalisation. Networks should adapt to the user and not *vice versa.*

Stakeholder groups of clinicians, educationalists, or computer scientists require mission statements expanding to cover: aims, objectives, test techniques and targets. The websites should support audit of key topics with templates, and a structured testing programme.

It is necessary to decide the degree of visibility to users, other stakeholders and indeed the rest of the world. To take the work forward, steering committees may serve to improve implementation. Although the aim is to minimize the effort of administration of the future surgical educational systems, steering groups will be needed to get the system running in the first phase. The following groups are considered:

- New Function Review Group (NFRG) – Appendix 9.2
- Standards Review Group (SRG) – Appendix 9.3
- New Technology Review Group (NTRG) – Appendix 9.4

Compliance with Academy of the Royal Colleges policies upon performance and also HST Guidelines is essential. Each review panel will set its own review priorities, and initially elect who is on the panel, and how often to meet, most administration being handled electronically.

Integrated policy formation in education: Incorporating computer-based learning within the existing educational infrastructure sets demands upon establishing a computer-based surgical simulation infrastructure. It will be tempered by the pragmatic solutions found using an integrated approach between regulatory authorities, research institutes and industry. In particular, this relates to advances in new surgical tool technologies. Surgical simulation is still early in its development. Without a rational integrated policy, though, it is likely to be under-resourced and under-utilized. The author advocates a coherent research and development strategy bringing technologies and expertise from the different disciplines together. Simple and reliable evaluation is key to the integration of these new resources into the established training infrastructures. The DRJ has demonstrated a way forward.

9.12 Going global

It has been said that the purpose of completing a PhD is to discover what your life's work will be. In the case of the development of virtual university systems, the tools are still in their infancy but a robust infrastructure has been established. The author believes the evidence supports the argument that the advancing media should be integrated into the working environment and the closer it gets to the actual 'reality' of the operating environment the better the training tools will be. There are however various topics that need to be developed further. Seeing beyond the horizon is going to require the collaboration of many different groups in this multidisciplinary approach. The way forward could be through the consortium of the support of Virtual Orthopaedic University by the *World Universities Network (WUN).*

This is a grouping of sixteen research-led institutions of international standing who have come together to create a worldwide research and graduate education partnership. The network aims to meet the challenges of research in rapidly developing, often interdisciplinary areas, which is of global significance. It will be necessary to draw upon the increased capability, diverse expertise, creativity and the wide range of techniques that come from international collaboration. This exemplifies the interrelated basic science domains. Experienced researchers in these fields are scarce, widely distributed and highly specialized. There is significant value from bringing these complementary foci across institutions and nations together.

The goal is to develop an effective framework that provides coherence and a supportive organizational environment, utilizing innovative methods of working and emphasizing the rapidly developing web-based technologies, to effectively support long-term (and long-distance) interactions and relationships. The *VOU framework* can support this enhancing the role of simulation within the interface for the different clinical and research domains. By facilitating faculty coming together on an international basis and providing intellectual venture capital and supporting infrastructure, WUN-VOU will reduce the barriers to international collaboration, makes links less fragile and more likely to grow over time, creating the opportunity to undertake research of a scale and scope that is beyond that of individual institutions.

The next stage will be to establish an initiative to integrate the pathway *'from concept to operating theatre'* since the relevant basic sciences are already on board. These include:

- Informatics (VOU development itself, including the educational aspects and course preparation).
- Bio-medical Informatics (e.g. the DRJ).
- Stem Cells and Tissue Engineering (primarily bone and cartilage).
- Nano-technology (especially interfaces).
- Advanced Materials (often ceramic).
- Mobile Devices and Wireless Communications overseeing the implementation (extending the work of the equator project) (165).
- Public Policy and Management (Building the trust and security services)

The network's teams of faculty from centres of excellence have the collective expertise to tackle major global challenges and this is just the initiative required during the WHO 'Decade of the Bone and Joint', whether on behalf of corporate partners or global agencies. This can help to develop research proposals for public funding, in ways that are not feasible by individual institutions.

9.12.1 Registration of domains

VOEU derivatives need to be registered within University domains to allow access via the Joint Academic Network, and as such would require registration of its name[32]. This may be an important issue since registering a 'new' university is a major undertaking, requiring review by curriculum committees from each country, and is practically an unrealistic logistical exercise. It is also likely to raise many political concerns. It may therefore be necessary to register the name elsewhere or to change it. This matter needs to be resolved by the end of the beta-testing phase before the planning of Version 4 multicentre trials.

9.12.2 Web identity

The URL that individuals will use for the access to VOEU needs to be agreed. This will influence the *'Educational Identity'* of future surgical educational systems. This is likely to be based upon the structure: http://vu.subject.region.academic.country

[32] An example of this is http://vou.soton.ac.uk.

Possible options are: *.org, .eu.edu, .ac.uk*. Currently the beta-testing demonstrator for orthopaedics is at: www.vou.soton.ac.uk .

9.12.3 Projected developments

Over the coming year or two, as part of the ongoing development of future surgical educational systems, the policy is likely to develop where the emphasis is to delegate the effort of other multimedia educational orthopaedic module (MEOM) generation. The following protocols are agreed: IMS, OSML keywords, JCHST Performance, and so the questions are what VOEU offers the educational establishments such as the RCS:

- Generic system for administration of courses
- Extension to basic surgical training flexible and distance learning infrastructure
- Feedback upon surgical training and practice performance
- Educational research infrastructure

What future surgical educational systems need from the RCS and other equivalent establishments are:

- Dynamic review journal content
- Surgical review board support
- Style guide (for BST / HST integration)
- Orthopaedic course structure templates
- Media best practice guide

9.12.4 Components from institutions

Course core modules will need to be included and the addition of list servers and teleconferencing archives may help to enhance this. Access to the main specialty journals is essential, and specific support for master classes should include uploading of additional lectures, especially guest lectures. The presentation of conference proceedings is anticipated. Redefinition of the UI Design and navigating through the simulations as a route to relevant content using adaptive hypermedia links are seen as a way of changing the user's perspective. In this way, the textual information follows the visual reference rather than vice versa, as the simulation environments are developed further.

Human Computer Interface considerations include the 'traditional' - Windows, Icons, Menus, Pointers (WIMP) *etc.* to be agreed and also submission of the design overview for review by Education Groups responsible for the distribution of developer templates. This is why a cascading style sheet *(.css)* approach was adopted.

9.12.5 Wider support for the project

A clear and concise plan for the longer-term funding approach of VOEU derivatives needs to be agreed by all stakeholder groups. To assist with the implementation of the above plans, it is essential to mobilize support in all groups, whose interests coincide with future surgical educational systems. This is initially to obtain material of a quality suitable for presentation in future surgical educational systems, and ultimately to seed the process for the next generation of the platform.

This can be assisted by developing a list of website addresses for contact. Awareness has been generated by the user of list server functions for the mailing, informing interested parties through the distribution of presentation packs (including the *ICT in healthcare* text). In this way experts such as Prof. Bulstrode (Oxford) and Mr Oliver (Edinburgh) in 2000 have been introduced to the concepts.

9.13 Conclusions

9.13.1 Embedding surgical simulation

In the spirit of future surgical educational systems, this thesis has recorded the present state of the evolution of educational strategy and so the infrastructure proposal is now fixed for initial periods using simulation technologies. The technologies provide a reusable, upgradeable and cost-effective resource for the healthcare of all. This promotes compatibility and integration of evolving digital surgical technologies. The aim remains to ensure adequate standards of realism in modelling.

By developing the surgical simulation system and embedding it into the dedicated educational environment it has been possible to demonstrate the positive value of integrated learning environments. These are only going to be effective if the end-user can both generate and review the content in a structured orderly fashion with minimal lead time and investment of effort to create new high-quality material. The purpose of the evaluation was to confirm that the design philosophy is on track and making adequate progress, which it is. This involves the establishment of new techniques for the collection, distribution and presentation of information in video form.

Users can now build libraries that are truly integrated multimedia. The need to develop more advanced media is driven by the need for clinicians' continuing professional development. This will require a flexible *'virtual university infrastructure'* that is pervasive and mobile. It will need the robust peer review process to build confidence in the system, as individuals must trust material collected for a common goal from multiple sources.

9.13.2 Assessment of surgical trainees – the staff

The other side of the two-edged sword, collecting data for surgical skills performance will eventually provide a quantitative performance marker for the surgeon, the critical factor is where these standards need to be set. This is a political as well as ethical dilemma to be met head on. Should the policies outlined in the previous sections be implemented, progress to a new era of quantitative as well as qualitative assessment of surgical performance will arrive. The current climate of surgical practice may best be described as emerging – as Figure 9.3 demonstrates;

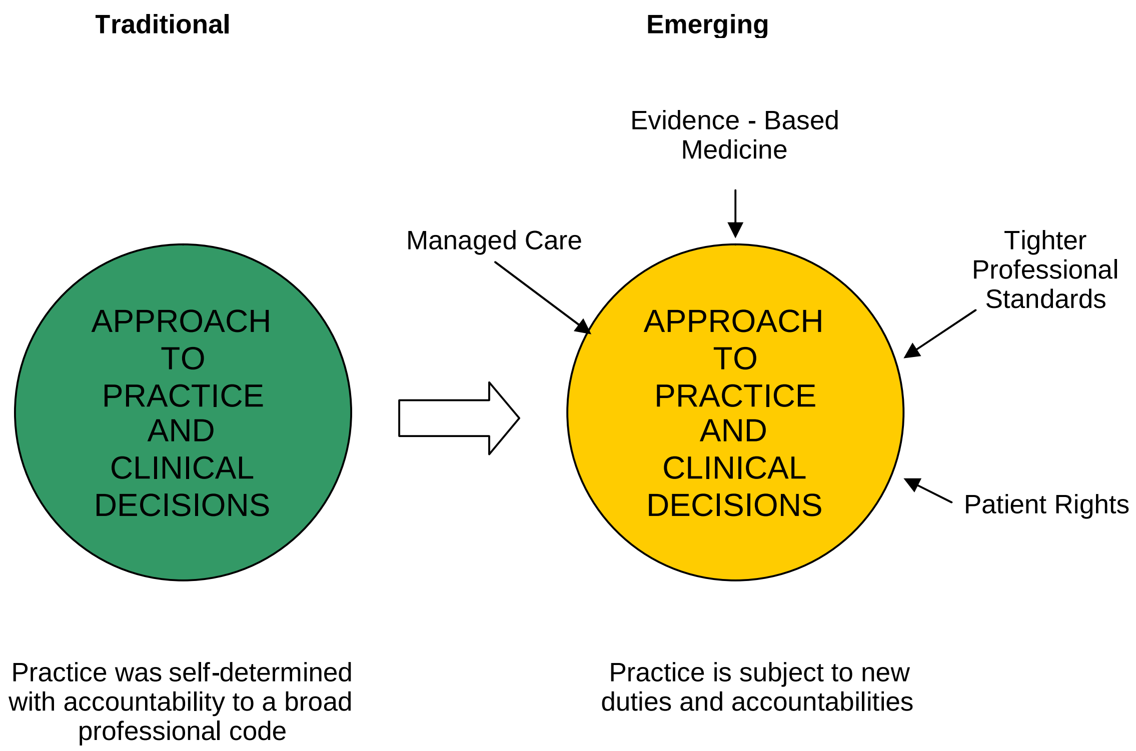


Figure 9.3 Clinical autonomy: an evolving concept (After LJ Donaldson)

9.13.3 Conclusions from the 1st & 2nd prototype simulations

Capturing useful video worlds is possible, but further development of the capture system is likely to require a degree of automation. The original simulation is run at 20 frames / second, thus the illusion of real time playback is possible. Using ATM + the K-Net™ video *codec* increased throughput 10 fold, but the limited access to the system and high maintenance led to a change of approach. Future work will involve developing the technique in the clinical environment and the development of integrated computer graphics for the development of a more interactive system with pathology incorporated. It is essential that such a technology finds a niche within the rapidly developing framework of the computer-based (distance) learning (CBL) strategies of the Medical Educational Establishment, and thus development alongside the current medical educational infrastructure reforms will be necessary.

With the development of photo-rendering of graphical simulators in real time, it is possible to have a system that can be effectively used to teach pattern recognition to surgical trainees, whilst allowing them the opportunity to manipulate the virtual environment - which can have other qualities such as haptic feedback. Since this is not yet a reality, the system proposed for the development of a purely video based system allows for a starting point to obtain adequate video data to incorporate into the photo-rendering process.

These parameters are likely to be upgraded again and again over the next few years in this interim period, as disc storage becomes both cheaper and less significant, ultimately allowing for disc-less systems. Photo-rendering means that if one attempts to develop the field further, it will be necessary to progress toward larger databases that can be collected in a more automated fashion and explore parallel image transformation algorithms such as exploiting salient stills technology, and 'morphing' from key frames.

Currently this system is faster and cheaper than graphical based systems, producing images of a higher fidelity, albeit with limited interactivity. It has the potential to be integrated with graphical systems to greatly enhance their performance, and in the future could have a significant role to play in texture rendering in graphics worlds. The AIPES measure of simulator performance will progressively improve in accuracy and ultimately provide a system of assessment with the underlying foundation of a large population of surgeons, due to the use of distributed computing architecture to develop this important database.

9.13.4 Original approach to simulation

Developing an original approach to surgical simulation involves accurately modelling the specific skills and requirements of surgeons. These in effect outline the human side of the HCI equation. The human performance aspects of this have far-ranging implications for the potential way that surgeons are screened in future for their training. Education dictates technology, and the real issue of whether VR is desirable is not easily answered. It is certainly desirable for the training of procedural skills and this has already been borne out in the other domains of vehicle and high-risk (*e.g.* nuclear) industries.

The thesis supports the hypothesis that *'a virtual university infrastructure offers the potential of providing an effective distributed learning environment for orthopaedic surgeons'* by demonstrating that the integration of the user, the educational scenario and the simulation supports a successful outcome. This should apply to surgical training as qualified time, in the same way as flight simulators, for it offers potentially significant cost-effectiveness. The surgical VR implications are psychological, social and medical. This is not only desirable, but also motivated by concern over assurance of adequate levels of training.

9.14 Final word

This thesis has demonstrated that users have a significant role to play in the design cycle and in the integration of key components. The book *Information and Communication Technology (ICT) Skills for healthcare Professionals* was written to support the necessary cultural transition, along with the development of the Dynamic Review Journal to facilitate continuous review of data to ensure that it is updated in line with new analyses and is scalable, using a client-server architecture, aiming for low maintenance to ensure higher compliance from an end-user population who need critical systems support. They have been previously disappointed by unrealistic claims.

Further enhancement of the cultural change support comes from the personal profiling so the users sense that the system is 'theirs' and that they have a significant role to play in its use and management. Further development requires strict feasibility studies with thorough cost-time calculations. Other issues to address are cross-border searching (Data Protection Act), making the portal 'intelligent', and responding to the user characteristics – 'give me what I want, not necessarily what I asked for'. This naturally raises funding issues and the author has proposed short, medium, and long-term solutions to this.

It is not possible to include everything in a simulation and thus it is worth considering the initial assumption that the simulation is essential for teaching certain tasks. It is certainly the key to teaching some tasks but only those to which it is most suited, *i.e.* procedural training. The multicentric training needs are often driven by external factors, such as specific courses or adaptation for local populations or preferences, such as protocols.

Variables may ultimately be used for *'fitness to practice'* measurements and this needs to be incorporated into the training process. They may not relate to competence, or to error rate. They may include the complication rates and the time for a given interaction.

It is dubious that judgement should be based solely upon tools since they do not include other aspects such as communication. A fundamental framework for the constraints that this system has to obtain is now available. The use to which measurements are put is driven by external factors. This includes interaction the multicentric user needs and measurements of the effectiveness can be gleaned from a wider population study. Ultimately, whatever the technology it is likely that the attitude of the user will have a pivotal role to play in the effectiveness of the educational process.

Fundamental to the uptake of the system is peer review, and this has been built into the design with updating and rejection of submissions with appropriate reasoning. The author therefore can recommend the system to my orthopaedic colleagues, convinced that this is a progression, representing a digital conversion for the paper-based processes, reflecting evolution not revolution. Although initially perceived as a 'Brave New World' by some, it is in fact the natural progress from one medium to another and really represents 'Better, Faster, Cheaper' approaches to the age-old problems. Surgical systems of the future will need to augment the humanity and reality of surgery, described as the eyes of a hawk, the heart of a lion and the hands of a lady, with the mind of a machine.

10 APPENDICES

10.1 Appendix 2.1 ICT Skills for Healthcare Professionals

A copy of this book is attached in the pocket in the back cover of this volume. Further copies are available upon request: simon@grange.me.uk

10.2 Appendix 3.1 Frame density estimations

Device for recording	Granularity (Resolution)	Frames per cc, in 3D	Allowing for frames recorded every 5 degrees of rotation	Allowing for frames recorded every 3 degrees of rotation	Allowing for frames recorded every 1 degree of rotation
The Bird™	2mm	5 x 5 x 5 = **125**	125 x 72 = **9,000**	125 x 120 = **15,000**	125 x 360 = **45,000**
Optitrack™	1mm (25dpi)	10 x 10 x 10 = **1,000**	1,000 x 72 = **72,000**	1,000 x 120 = **120,000**	1,000 x 360 = **360,000**
	0.5mm (50dpi)	20 x 20 x 20 = **8,000**	8,000 x 72 = $\mathbf{5.76x10^5}$	8,000 x 120 = $\mathbf{9.6x10^5}$	8,000 x 360 = $\mathbf{2.88x10^6}$
	0.3mm (75dpi *i.e.* > normal visual acuity)	33 x 33 x 33 = **36,963**	36,963 x 72 = $\mathbf{2.66x10^6}$	36,963 x 120 = $\mathbf{4.43\ x10^6}$	36,963 x 360 = $\mathbf{1.33x10^7}$

Appendix 3-1Frame density estimations

10.3 Appendix 3.2 Joint capacity relationship to hard disc space

Resolution (x,y,z)[mm]	**2**	**1**	**0.5**	**0.33**	**2**	**1**	**0.5**	**0.33**	**2**	**1**	**0.5**	**0.33**
Resolution (α)[degrees]	5°	5°	5°	5°	3°	3°	3°	3°	1°	1°	1°	1°
Frames per cc	9×10^{3}	7.2×10^{4}	5.76×10^{5}	2.66×10^{6}	1.5×10^{4}	1.2×10^{5}	9.6×10^{5}	4.43×10^{6}	4.5×10^{4}	3.6×10^{5}	2.88×10^{6}	1.33×10^{7}
5cc	4.5×10^{-1}	3.6	2.88×10^{2}	1.33×10^{3}	7.5×10^{0}	6.0×10^{0}	4.8×10^{2}	2.22×10^{4}	2.25×10^{1}	1.8×10^{2}	1.44×10^{3}	6.65×10^{4}
10cc	9×10^{-1}	7.2	5.76×10^{2}	2.66×10^{3}	1.5×10^{1}	1.2×10^{1}	9.6×10^{2}	4.44×10^{4}	4.5×10^{1}	3.6×10^{2}	2.88×10^{3}	1.33×10^{5}
15cc	1.35	1.08×10^{1}	8.64×10^{2}	3.99×10^{3}	2.25×10^{1}	1.8×10^{1}	1.44×10^{3}	6.66×10^{4}	6.75×10^{1}	5.4×10^{2}	4.32×10^{3}	2.66×10^{5}
20cc	1.8	1.44×10^{1}	1.15×10^{3}	5.32×10^{3}	3.0×10^{1}	2.4×10^{1}	1.92×10^{3}	8.88×10^{4}	9.0×10^{1}	7.2×10^{2}	5.76×10^{3}	3.99×10^{5}
25cc	2.25	1.8×10^{1}	1.44×10^{3}	6.65×10^{3}	3.75×10^{1}	3.0×10^{1}	2.4×10^{3}	1.11×10^{5}	1.1	9.0×10^{2}	7.2×10^{3}	5.32×10^{5}
30cc	2.7	2.16×10^{1}	1.73×10^{3}	7.98×10^{3}	4.5×10^{1}	3.6×10^{1}	2.88×10^{3}	1.33×10^{5}	1.35×10^{2}	1.08×10^{3}	8.64×10^{3}	6.65×10^{5}
35cc	3.15	2.52×10^{1}	2.02×10^{3}	9.31×10^{3}	5.25×10^{1}	4.2×10^{1}	3.36×10^{3}	1.55×10^{5}	1.58×10^{2}	1.26×10^{3}	1.01×10^{4}	7.98×10^{5}
40cc	3.6	2.88×10^{1}	2.3×10^{3}	1.06×10^{4}	6.0×10^{1}	4.8×10^{1}	3.84×10^{3}	1.78×10^{5}	1.8×10^{2}	1.44×10^{3}	1.15×10^{4}	9.31×10^{5}

Table 3-2 Joint capacity relationship to hard disc space

10.4 Appendix 4.1 AIPES scoring system version 2.01

Virtual Environment - Autonomy, Interactivity & Presence
The objects in the computer-generated world behaved in a natural way?
The objects in the computer-generated world responded realistically to your inputs?
Your interaction with the virtual objects was realistic?
The feedback provided by the hand controller was realistic?
You experienced a sense of being "really there" inside the virtual environment?
The objects in the computer-generated world looked real?
The computer generated world seems to be something that you saw (low score) or something that you visited (high score)
Additional comments about new developments of surgical presence (sense of immersion in a virtual environment) on the horizon and how these might be integrated?..
Additional comments about what different types of man-machine interface you are familiar with, e.g. games boxes and devices, and what they portray to you, the user. ....
Environment - the measure to which the user feels that they are in an appropriate environment for learning.
Simulating specific events is possible so that there is no risk to the patient
The simulation allows reproduction of a specific event repetitively
The simulator allows control of independent multiple variables
The environment provided Ease of Access, comfort, and did not impose unnecessary time penalties
The Environment was adequate for the task - not distracting from the effort
The simulation environment is appropriate e.g. a screen displaying image guided surgical simulation

Additional comments about planning to develop broad bandwidth Information Technology infrastructures to provide a uniform quality of service for internet based users (Do or are you planning to use ADSL)...
Scenario - the educational tool for the user.
The system adequately provides an opportunity to hold simulation for discussion and teaching
Allows errors to be made and explanation of their consequences given
The simulator provides an adequate record and critique of performance
The simulator allows objective evaluation of performance
The simulator allows uncommon events to be experienced
Additional comments on what the consequences of inadequacies in the simulated environment are, and the likely error in the real environment, leading to entry-level standards for individual sub-skills training in simulation..
Additional comments on how these new criteria can be integrated into a coherent set of guidelines for the introduction of these new technologies into the operating and training environments within the existing European Health Care infrastructures ....
Additional comments on what simulators should be used for, i.e. Which surgical sub-skills are being assessed? ....
Neurolinguistic programming (NLP) - How we think 1 = strongly disagree, 7 = strongly agree
When operating in a Minimal Access Surgical Environment, I believe that my perception of reality is predominantly dependent upon my Vision
When operating in a Minimal Access Surgical Environment, I believe that my perception of reality is predominantly dependent upon my Auditory sensitivity
When operating in a Minimal Access Surgical Environment, I believe that my perception of reality is predominantly dependent upon my Kinaesthetic sensitivity

Additional comments about how easy you felt the software was to become familiar with: - ....
Adopting the NASA TLS scale for the evaluation of the simulation experience. This requires the user to rate the value of the simulation in a more pragmatic level of communication. These reflect mental, physical, & temporal demand, performance, and effort.
Mental Demand - How much mental and perceptual activity was required (e.g. thinking, deciding, remembering)
Mental Demand - Was the task easy (low score) or complex (high score)
Physical Demand - How much physical and perceptual activity was required (e.g. turning, controlling, activating)
Physical Demand - Was the task restful (low score) or laborious (high score)
Temporal Demand - How much time pressure did you feel at the rate of pace of the task elements occurred.
Temporal Demand - Was the pace slow (low score) or rapid (high score)
Performance - How successful did you feel that you were in completing the set tasks?
Performance - How satisfied were you with your performance accomplishing these goals? Unsatisfied (low score) or very unsatisfied (high score)
Effort - How hard did you have to work (i.e. mentally and physically)
Effort - Was the effort low (low score) or high (high score) to accomplish your level of performance
Any other addition comments you wish to make?

The questions that are most discriminatory question for each aspect of the review will be kept for future trials, based upon the widest range of responses.

10.5 Appendix 7.1 Neurolinguistic programming (NLP)

As part of the process of developing other Pilot Study VALIDATION Methods in the 1st prototype simulation, to reflect what really matters to the surgeon, participants were asked Neurolinguistic programming questions to see if a relative balance of the sensory modalities could be obtained. Scores were in the range from 1 – 7 *i.e.* between ‘strongly disagree’ and ‘strongly agree’. The following questions (in Appendix Table 7.1) were posed to the participants:

Question	**Mean**
When operating in a Minimal Access Surgical Environment, I believe that my perception of reality is predominantly dependent upon my Vision	5.95
When operating in a Minimal Access Surgical Environment, I believe that my perception of reality is predominantly dependent upon my Auditory sensitivity	2.10
When operating in a Minimal Access Surgical Environment, I believe that my perception of reality is predominantly dependent upon my Kinaesthetic sensitivity	4.90

Table 7.1 Neuro-linguistic programming (NLP) questions in the EVW pilot study

These were adjusted to base 10 in order to make the results more manageable.

	Vision	**Auditory**	**Kinaesthetic**
Average Score (max. 10)	8.37	2.45	6.99
% of total	47	14	39

Table 7.2 Neuro-linguistic programming EVW pilot study results (%)

Multiplying by 10/7 to give a decimal value that was then described as a percentage to relate the relative importance of each factor. Clearly different aspects of a simulated environment reach human consciousness at different times, since this is in effect a single channel, although Multichannel, multitasking will be being performed within the neural networks on a sub-conscious level. Significantly the figures are relatively consistent implying that the opinion of the surgeons is uniform, in nature.

When operating in a Minimal Access Surgical Environment, *Table 7.2.3* details the Neurolinguistic Programming Analysis.

	Question	Variable Short Name	Variable Long Name NLP_	Type Range 1 - 7	Min	Max	Mean
V	I believe that my perception of reality is predominantly dependent upon my Vision	Vis	Visual	Integer (radio button)			
A	I believe that my perception of reality is predominantly dependent upon my Auditory sensitivity	Aud	Auditory	Integer (radio button)			
K	I believe that my perception of reality is predominantly dependent upon my Kinaesthetic sensitivity	Kin	Kinaesthetic	Integer (radio button)			

Table 7.3 Neurolinguistic programming analysis

10.6 Appendix 8.1 User profile

Experience	Range of Values	Mean
Years In Specialty	0 - 22	4.3
Confidence at shoulder arthroscopy	1 – 7 (TLS highest and lowest scores)	2.55

Appendix Table 8.1 User clinical experience results in the 3rd design phase simulation

	BST	HST1,2	HST3,4	HST5,6	CS	Tutor	Researcher	Other
Number	8	1	3	1	0	0	0	0
%	61	8	23	8	0	0	0	0

Appendix Table 8.2 User occupational status (stakeholder class)

N=19	24 or younger	25-34	35-44	45-54	55 or older
Number	1	15	3	0	0
%	5.3	79	16	0	0

Appendix Table 8.3 User age at time of the 3rd generation simulation trial

N=19	Minimum	Maximum	Mean
Number	0.9	16.0	4.94

Appendix Table 8.4Length of service to date

N=19	Daily	Weekly	Once a Month	Rarely	Never
Number	11	5	1	0	2
%	57.9	26.3	5.3	0	10.5

Appendix Table 8.5 Experience of using the Internet/World Wide Web: Use of the Web at home

N=19	Daily	Weekly	Once a Month	Rarely	Never
Number	14	4	0	1	0
%	73.7	21.1	0	5.3	0

Appendix Table 8.6 Frequency of web access at work (not just for work related information)

Do users find the:	Agree (%)	Disagree (%)	Neither (%)
High-tech route to information is intimidating	4 (21.1)	13 (68.4)	2 (10.5)
Paper-based system easier to use	6 (31.6)	8 (42.1)	5 (26.3)
Paper-based system a more effective means of sharing information	3 (15.8)	12 (63.2)	4 (21.1)

Appendix Table 8.7 General experience with web-based technologies

Why do you use the Network/Web?	Agree (%)	Disagree (%)	Neither (%)
It is the only method available to retrieve the information	0 (0)	16 (84)	3 (15.8)
Convenience	16 (84)	1 (5.3)	2 (10.5)
Easier searching	15 (78.9)	1 (5.3)	3 (15.8)
Speed to retrieve information	15 (78.9)	0 (0)	4 (21.1)
Easier to use	15 (78.9)	1 (5.3)	3 (15.8)
Prefer to use	15 (78.9)	1 (5.3)	3 (15.8)

Appendix Table 8.8 Reasons for using networked technologies

User Grade	Number Performed in total	1st design phase simulation study Number [N=21] (%)	3rd generation simulation study Number [N=13] (%)
Senior House Officers	None	0 (0)	8 (62)
	Less than 10	6 (29)	0 (0)
	Between 10 and 50	0 (0)	1 (8)
Specialist Registrars and Career Grade Surgeons	None	0 (0)	0 (0)
	Less than 10	9 (43)	2 (15)
	Between 10 and 50	4 (19)	2 (15)
Consultants	Greater than 50	2 (9)	0 (0)

Appendix Table 8.9 Number of arthroscopies performed

	Question	Min	Max	Mean
T	Trainer	0	0	0
P	Performed	0	2	0.1
S	Performed Supervised (Scrubbed)	0	3	0.2
S	Performed Supervised (Unscrubbed)	0	3	0.2
A	Assisted	0	5	0.5

Appendix Table 8.10 Experience of this type of surgery and in what capacity

	Question	Range	Min	Max	Mean
1	Did the objects in the computer-generated world behave in a natural way?	1 - 10	4	10	7.4
2.	**Did the objects in the computer-generated world respond realistically to your inputs?**	**1 - 10**	**4**	**10**	**7.4**

Appendix Table 8.11 Analysis of the autonomy of the 3rd generation simulation (simulation) system

	Question	Range	Min	Max	Mean
3.	How realistic was your interaction with the virtual objects?	1 - 10	4	10	6.9
4	**Was the feedback provided by the hand controller realistic?**	**1 - 10**	**1**	**10**	**6.1**

Appendix Table 8.12 Analysis of the interactivity of the 3rd generation simulation (simulation) system

	Question	Range	Min	Max	Mean
5	To what extent did you experience a sense of being "really there" inside the virtual environment?	1 - 10	1	10	5.5
6.	Did the objects in the computer-generated world look real?	1 - 10	3	10	7.0
7.	**The computer-generated world seems to be something that you saw (low score) or something that you visited (high score)**	**1 - 10**	**1**	**10**	**6.0**

Appendix Table 8.13 Analysis of the presence of the 3rd generation simulation (simulation) system

	Question	Variable Short Name	Variable Long Name Environment	Type Range 1 - 10	Min	Max	Mean
8.	Simulating specific events is possible so that there is no risk to the patient	Env1	risk1	Integer (radio button)	2	10	6.7
9.	The simulation allows reproduction of a specific event repetitively	Env2	events2	Integer (radio button)	3	10	6.7
10.	The simulator allows control of independent multiple variables	Env3	imv3	Integer (radio button)	3	9	6.6
11.	The environment provided Ease of Access, comfort, and did not impose unnecessary time penalties	Env4	access4	Integer (radio button)	4	10	7.1
12.	The Environment was adequate for the task – not distracting from the effort	Env5	distract5	Integer (radio button)	2	10	6.5
13.	The simulation environment is appropriate *e.g.* a screen displaying image guided surgical simulation	Env6	igs6	Integer (radio button)	3	10	6.7

Appendix Table 8.14 Environment of the 3rd generation simulation (simulation) system

	Question	Variable Long Name Education	Type Range 1 - 10	Min	Max	Mean
14.	The system adequately provides an opportunity to hold simulation for discussion and teaching	discussion1	Integer (radio button)	2	10	6.2
15.	Allows errors to be made and explanation of their consequences given	errors2	Integer (radio button)	3	10	6.6
16.	The simulator provides a record and critique of performance	record3	Integer (radio button)	2	10	6.1
17.	The simulator allows objective evaluation of performance	evaluation4	Integer (radio button)	2	10	6.1
18.	The simulator allows uncommon events to be experienced	events5	Integer (radio button)	4	10	7.1

Appendix Table 8.15 Educational scenario evaluation

	Header		Range	Min	Max	Mean
22.	Mental Demand - How much mental and perceptual activity was required (e.g. thinking, deciding, remembering)	TLS1	1 - 7	2	7	4.35
23.	**Mental Demand - Was the task easy (low score) or complex (high score)**	**TLS2**	**1 – 7**	**1**	**6**	**3.76**
24.	Physical Demand - How much physical and perceptual activity was required (e.g. turning, controlling, activating)	TLS3	1 – 7	1	4	2.47
25.	**Physical Demand - Was the task restful (low score) or laborious (high score)**	**TLS4**	**1 – 7**	**2**	**5**	**2.94**
26.	Temporal Demand - How much time pressure did you feel at the rate of pace of the task elements occurred.	TLS5	1 – 7	1	5	3.24
26.	**Temporal Demand - Was the pace slow (low score) or rapid (high score)**	**TLS6**	**1 – 7**	**3**	**6**	**4.12**
26.	Performance - How successful did you feel that you were in completing the set tasks?	TLS7	1 – 7	2	6	4.41
26.	Performance - How satisfied were you with your performance accomplishing these goals? Unsatisfied (low score) or very unsatisfied (high score)	TLS8	1 - 7	1	6	4.06
26.	Effort - How hard did you have to work (i.e. mentally and physically)	TLS9	1 - 7	2	6	3.65
26.	Effort - Was the effort low (low score) or high (high score) to accomplish your level of performance	TLS10	1 - 7	2	5	3.65

Appendix Table 8.16 Analysis NASA TLS

10.7 Appendix 8.2 VOEU simulation trials protocol v2.01

This protocol for the evaluation of the surgical simulation systems is based around the following checklist to be completed as part of a usability programme for new users being introduced to the system. It consists of the following steps;

1. Introduction

Users are presented with a page of introductory text outlining the role of the system, the process of testing, and the research functions that are proposed to benefit from the use of results

2. Demographic Questionnaire

This is subdivided into various topics. These are;

User Clinical Experience (Questionnaire A)

User ICT Literacy (Questionnaire A)

Usability analysis of the integrated simulation environment (Questionnaire C)

3. Opportunity to familiarize with the simulation system

Users are given a brief demonstration of the system to see how the components work and are then invited to take over the controls.

4. Simulation task performance

User simulation perceptions are gauged by allowing 2 - 3 minutes of `hands on' experience with the simulator for users to familiarise themselves with the system, which was using a standard web browser and mouse for the user interface, as compared to the conventional surgical equipment. Users were subjected to an exercise, which involved assessing the simulation itself. The users were asked to negotiate the route through the shoulder joint that would represent the standard navigation conducted as part of a diagnostic shoulder arthroscopy. The order (sequence of steps) of this was based upon the standard arthroscopy as outlined by Mr Bunker (Consultant Shoulder Surgeon)(1).

5. Test performance

Once they are familiar with the interface (to their satisfaction), they are then given a set task to perform according to a formal question provided in the VOEU *`virtual classroom'* environment. They are then requested to complete the AIPES scoring system for surgical simulation (Questionnaire C). These perceptions are dependent upon the time and context of the study. With a rapidly evolving world of media that exposes the users to expectations, which are much, higher for image quality and with convenient interfaces it is important for users to appreciate that this is a research trial, and not a system prepared for commercial exploitation.

This involves a time-limited exercise to enter the virtual environment (VE). This could for example require the user who is navigating a shoulder arthroscopy simulation to obtain a good view of three structures inside the joint, e.g.:

- o Insertion of Long Head of Biceps
- o Foramen of Weitbrecht
- o Inferior Recess

Another surgeon would confirm identification independently; else the pages recalled by the simulator confirm the location of the user. All surgeons should be able to perform the task within reasonable time (less than 60 seconds).

Perception [N=17]	**Range Minimum / Maximum**	**Mean The 1st design phase simulation**	**Mean the 3rd design phase simulation**
Computer Assessment	2 - 6	5.38	
Vision	2 – 6	5.95	4.71
Auditory	1 - 6	2.10	2.82
Kinaesthetic	3 - 7	4.90	5.12

Appendix Table 8.17 Attitudes of the user population to quality of the simulation

10.8 Appendix 9.1 Principles of confidentiality

These rules are non-negotiable. Each country has its own mechanisms in place to assure protection of data stored regarding the rights of individuals. The data protection act in the UK must be adhered to. Transferring data from a clinical investigation (e.g. MRI scanning) for the preparation of training material requires a process of anonymisation. For the reasons indicated below, it is both necessary to ensure that this data is normally not traceable to source except by a recorded responsible third party such as the primary care doctor who may need to identify the patient if new information comes to light.

The patients consent for their data, and information derived from it, must be obtained. For the purposes of VOEU, and any derivatives, this must include a signed document in accordance with the laws of the country where the patient is being treated. For the purposes of future surgical education systems, it is duty of the attending surgeon to obtain the permission of the patient for the material to be used for educational purposes including presentation if it can be traceable, and managed using the system for non-repudiation. The requirements are as follows;

1. Combinations of individuals or groups may have access to parts or all of a document body.
2. Combinations of individuals or groups may have access to parts or all of a document header.
3. Entitlement access restrictions to part of a document may be overridden for emergent or critical care situations, or with documented patient authorization.
4. Access entitlement can change over time and context with appropriate permissions.
5. Access entitlement restrictions can be added by healthcare providers or by patients.
6. Entitlement restriction removal and entitlement granting can only be by documented patient authorization.
7. Access entitlement and entitlement restrictions must travel with the document transmission. This may be via a linkbase.

10.9 Appendix 9.2 New Function Review Group (NFRG)

Any new function that is to be included within the virtual university should be approved by the NFRG. This is to ensure that the functions do not conflict in nomenclature – *i.e.* by the integration within the XML schemas etc. The following four responsibilities should be met: technical considerations, surgical content considerations, ergonomics, publishing, and dissemination.

The preparation of the new functions will necessitate the development of multiple components that will require authorisation from many particular Stakeholder groups. To simplify the processes can be illustrated, tracked as issues that are related arise, and displayed to emphasise paradigm shifts in thinking. Changes can then be recognised more easily and responded to. An example is the alerting system built into the discussion fora that can then be used for threaded discussion following up events and indeed used for adverse event reporting.

10.10 Appendix 9.3 Standards Review Group (SRG)

This group will be organised by the governing body. It will recognise and manage the content generation initially whilst the *Virtual University* is in its infancy. It must be able to organise systems to detect bias such as commercial interests. These may try to influence the syllabus, curriculum, or even the tutor directly. Certain schools are established leading to the *'chapel effect'*, the aim being not to create uniformity of the course rather to ensure baseline standards. The review group would be responsible for maintaining academic integrity. It should guide the research by overseeing the setting up the dynamic review process on a regional level this may be handled by the surgical training committee.

10.10.1Proposal for the dynamic review process

This will be the core responsibility of the SRG. Ultimately this will become the *Document Handling Process (DHP)* – This is designed for the use of all members of the future surgical educational systems, be it production of multimedia modules or the production internal documents. It is worth considering the work already done by the IEEE upon *Learning Object Metadata*. This underpins the VOEU metadata used in the file headers. The group will need to consider the relationship between the European Union efforts and the *Health Level 7* work in the USA, and also the potential to interface using *SCORM.*

10.11 Appendix 9.4 New Technology Review Group (NTRG)

This subgroup will decide upon which new technologies will be evaluated and accommodated. As services develop, it will be necessary to provide for ‘Integration of Services’ such as Teleconferencing / Telemedicine.

- A New Technology Evaluation and Assessment form – Linking with the input from VOU-Regional will be developed.
- Sub domains of the New Technologies awaiting reviewer selection, will be fed into the eventual syllabus of the Image Guided Orthopaedic Surgery (IGOS) modules. This group will oversee the interrelationship between the future groups of CAOS, *Computer Aided Surgery (CAS)* and *Computer Aided Minimally Invasive Surgery (CAMIS).*
- Confidentiality Standards for patient-specific data need to be signed up to by all users. This should be part of the automated registration process.

With the need for patient-specific data in the educational material, clearly we need protocols for peer review and anonymisation of data in order to ensure confidentiality.

11 Reference list

(In order cited)

(1) Osler W. The Principles and Practice of Medicine, 4 ed. Bayliss, Toronto, Canada 1999.

(2) Medical Protection Society. Can the Society Help Me Avoid the Possible Pitfalls of Practice. The Medical Protection Society Review, 6-7. 1998.

(3) Department of Health. A First Class Service: Quality in the New NHS. DOH; 1998. Report No: 98.

(4) Donabedian A. Evaluating the Quality of Medical Care. Millbank Memorial Fund Quarterly 4, 166-206. 1966.

(5) Edinburgh College Logbook Service. http://www.elogbook.org/ 17-2-2006.

(6) Sue Jackson SH. Improving organisational culture through innovative development programmes. International Journal of Health Care Quality Assurance 12[4], 143-148. 1-7-1999.

(7) European Computer Driving Licence www.ecdl.co.uk 17-2-0006.

(8) Gross M, Innovations in surgery. A proposal for phased clinical trials J Bone Joint Surg Br, May 1993; 75-B: 351 - 354.

(9) Jeppesen Instrument Pilot Manual. 1, JS314520, New York: Jeppesen Sanderson, Inc; 1999.

(10) Stotter A *et al*. Simulation in Surgical Training Using Freeze-dried Material. British Journal of Surgery 73, 52-54. 1986.

(11) Bevan PG. The Anastamosis Workshop. Royal College of Surgeons of England 63, 405-410. 1981.

(12) Reznick R. Object to Fine Skills Assessment: Bench Model Solutions. London: Royal College of Surgeons of England; 1999.

(13) Martin J *et al*. Objective Structured Assessment of Technical Skills (OSATS) for Surgical Residence. British Journal of Surgery 84, 273 -278. 1997.

(14) González M.A.Ballester AZ. Measurement of Brain Structures based on Statistical and Geometrical 3D Segmentation. Lecture Notes in Computer Science; Springer Verlag; 1998.

(15) Tony Trowbridge RH. Virtual Arthroscopic Knee Surgery Simulator. Robin Hollands 1996 January 1 Available from: URL: http://www.shef.ac.uk/uni/projects/vrmbg/arthro 1 .html

(16) McCloy R, Stone R. Science, medicine, and the future: Virtual reality in surgery. BMJ 2001 Oct 20;323 (7318):912-5.

(17) McCloy R. WM. A new method for laparoscopic surgery training. BJHC&IM 13[10], 29. 1996.

(18) Darzi A., Smith S., Taffinder N. Assessing Surgical Skill. BMJ [7188], 887-888. 1999.

(19) Darzi A. Future surgical robotic systems. RCS Fellows, RCS England 2003.

(20) Satava R M. The King is Dead. Interactive Technology and the New Paradigm for Health Care. Washington DC: IOS Press; 1995. p. 334-9.

(21) Medmark. 2004. http://www.shoulder.com/home.html

(22) Adam JA. Medical Electronics. IEE Spectrum 1994; 31(1):70-3.

(23) DiGioia AM *et al*. Medical Robotics and Computer Assisted Surgery in Orthopaedics: An Integrated Approach. 88-90. 1995. Washington DC, IOS Press. Interactive Technology and the New Paradign for Health Care.

(24) Megland R. Making Surgical Simulation Real. Computer Graphics [November], 37-39. 1996.

(25) Human Interface Technology Laboratory. Medicine and Virtual Reality, A Guide to the Literature. University of Washington State. 2001. http://www.hitl.washington.edu/proj ects/knowledgebase/medvr/medvr.html.

(26) Yasushi Yamauchi. Internet resources of computer aided surgery. 2001. http://www.aist.go.jp/NIBH/~b0673/englishcas.html.

(27) Westwood JD *et al*. Medicine Meets Virtual Reality 6. Studies in Health Technology and Informatics. Washington DC: IOS Press; Report No: 50. 1998.

(28) National Library of Medicine. 2001. Available from URL: http://www.nlm.nih.gov/pubsfactsheets/visiblehuman.html.

(29) UC Berkeley VESTA Lab Project. Available from URL: http://robotics.eecs.berkeley.edu/medical/ 2001.

(30) Driscoll M. Selecting the most appropriate WBT method. Web Based Training. 1 ed. San Francisco: Jossey-Bass Pfeiffer; p. 47-70. 1998.

(31) Bloom BS *et al*. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill; 1971.

(32) Spencer CF. Teaching and Measuring Surgical Techniques - The Technical Evaluation of Competence. Bulletin American College of Surgeons 63, 9-12. 1978.

(33) Kopta JA. An Approach to the Evaluation of Operative Skills Surgery. Am J Surg 70, 297-303. 1971.

(34) Smith R. Regulations of Doctors and the Bristol Enquiry. Need to be both Credible to both Public and Doctors. BMJ [317], 1539-1540. 1998.

(35) Bunker T.D. Shoulder Arthroscopy. In: Wallace W.A., editor. London: Dunitz; 1991. p. 9-23.

(36) Grange S., Bunker T., Cooper J. Networking virtual reality for shoulder arthroscopy. British Journal of Healthcare Computing 13[10], 26-28. 1996.

(37) Ota D., Loftin B., Saito T., Lea R., Keller J. Virtual reality in surgical education. Computers in Biology and Medicine 1995; 2(25):127-37.

(38) Satava RM. Virtual reality surgical simulator. The first steps. Surgical Endoscopy 1993; 3 (7):203-5.

(39) BOA Education Committee. Guide to core education for higher surgical training in trauma and orthopaedic surgery. London: British Orthopaedic Association; 1999.

(40) Orlansky J, Dahlman CJ, Hammon CP, Metzko J, Taylor HL, Youngblut C. The value of simulation for training. IDA paper. 1994.

(41) Rheingold H. Virtual Reality. 2 ed. London: Mandarin; 1991.

(42) Grange S. Chappel D. Trusting Digital Surgical Technologies. St. Andrews, Scotland, Strategic Issues In Healthcare Management Conference; 2002.

(43) McCloy R., Wilson M. A new method for laparoscopic surgery training. British Journal of Healthcare Computing & Information Management 13[10], 29. 1996.

(44) McCloy R, Stone R. Science, medicine, and the future: Virtual reality in surgery. BMJ 2001 Oct 20; 323(7318):912-5.

(45) Dubois E, Chavanon O, Cinquin Ph, Troccaz J. Analytical approach for the design of the surgeon's interaction with a computer assisted surgery system. Grenoble: Surgetica; 2002 p. 211-6.

(46) Raimbault M, Jannin P. Modeles de procedures chirurgicales pour la neurolochirurgie guidee par l'image. Grenoble: Surgetica; 2002 p. 2 17-23.

(47) Wargo M *et al*. Human Operator Response, Speed, Frequency and Flexibility. Washington DC: National Aeronautics and Space Administration; Report No.: NASA CR/874. 1987.

(48) Hartman BO. Psychology of Aerospace Medicine. In: Randall HW, editor. Aerospace Medicine. Baltimore: Williams and Wilkins; 1971.

(49) Coloughoun W. Variations in Mental Efficiency. Biological Rhythm and Human Performance. Academic Press; 1971. p. 39-107.

(50) Markes T *et al.* Intuitive Frequency Judgements as a Function of Higher Exceptions. Rice University: 70-76, 2001. http://cohesion.rice.edu/engineering/computerscience/tr/

(51) Goode AW. The Clinical Link in Surgical Competence. Challenges of Assessment in Training and Practice. London: Royal College of Surgeons of England; 1999.

(52) Phipps N, Gordon R, Wilkinson R. Dir; Thomas R. Doctor in the house. Rank 1954

(53) Mc Cloy R, Stone R. Science, medicine, and the future: Virtual reality in surgery. BMJ 2001 Oct 20; 323 (7318):912-5.

(54) Seymour NEM, Gallagher AGP, Roman SAM, O'Brien MKM, Bansal VKM, Andersen DKM, *et al.* Virtual Reality Training Improves Operating Room Performance: Results of a Randomized, Double-Blinded Study. Annals of Surgery 236[4], 45 8-464. 2002.

(55) N.Eldredge SJG. "Punctuated equilibria: an alternative to phyletic gradualism *i.* Models in Paleobiology. 2 ed. *Princeton, N.J.:* p. 82-115. 1985.

(56) S Grange. DRJ record of InSim3 trial. [Secure access required] http://www.voeu.ecs.soton.ac.uk/voeu/trials/cases.aspx?trial=55 2002.

(57) Ota D, Loftin B, Saito T, Lea R, Keller J. Virtual Reality in Surgical Education. Computers in Biology and Medicine 2[special], 127-137. 1994.

(58) Grange S *et al.* ICT Skills for Healthcare Professionals. Southampton University Press 2002.

(59) Grange S, Power G, Wills G, Bailey, Miles-Board T, Carr L, Hall W. Building a Virtual University into the Operating Theatre, Early Experiences of Implementation: p.56-63. VRIC; France 2004.

(60) Grange S. The AIPES scoring system for evaluation of surgical simulator performance: p.64-71.VRIC; France 2004.

(61) Bunker T.D., Schranz P.S. The art of diagnosis in the mystery shoulder. Clinical Challenges in Orthopaedics: The shoulder. 1st ed. Oxford: ISIS Medical Media; p.1-23. 1998.

(62) Cooper J., Ford L., Watson G. A Training potential for Non-invasive surgery. ftp://ftp.dcs.ex.ac.uk/pub/usr/lindsey/garth.ps.z, p.1-21. 1995.

(63) Philip Weaver, Nicholas Lambrou, Matthew Walkley. Practical SSADM. 2 ed. London: FT Pitman Publishing; 1998.

(64) Grange S, Bunker T, Cooper J. Virtual Reality, A training World for Shoulder Arthroscopy. UKVRSIG Conference Proceedings; Brunel University: Brunel University; 1997 p.159-68.

(65) Grange S, Bunker T. Surgery In The Digital Era. eBMJ [2nd November 2001]. 2001.

(66) Grange S., Bunker T. Surgical simulation using VR Technology The Current Issues. Annals of The Royal College of Surgeons of England [6], p.24-30 1998.

(67) Johnsen EG, Corliss WR. Human Factors Applications in Teleoperator Design and Operation. 1 ed. London: Wiley & Sons; p. 1-8. 1971.

(68) Witten WI. *et al.* Managing Complexity in a Distributed Digital Library. Computer 32[2], 74-79. 1999.

(69) Visarius H, Berlemann U. Concept and Clinical Aspects of Computer Assisted Spine Surgery. CAOS Computer Assisted Orthopaedic Surgery 1996.

(70) Nolte L-P, Berlemann U, Langlotz F, Wang Q, Jiang Z, Zamorano L. Image-Guided Insertion of Transpedicular Screws - A Laboratory Set-Up. Spine 1 995; 20 (4):497-500.

(71) Schwarzenbach O, Berlemann U. Jost B. Visarius H, Arm E, Langlotz F. The Accuracy of Computer Assisted Pedicle Screw Placement - An In Vivo CT Analysis. Spine. 1997.

(72) Grange S, Cooper J, Bunker T. Electro-optical Vs Electromagnetic Tracking Systems in Surgical Simulation. 1, 1-14. 2002. University Of Exeter, Department Of Computer Science.

(73) Brett PN. Touch Sensation in Minimal Access Surgery. Biomedical Engineering Research Institute 2004 Available from: URL: http://www.mee.aston.ac.uk/research/medical/touchmat.html

(74) Brett PN, Parkers TJ, Harrison AJ, Thomas TA, Carr A. Simulation of resistance forces acting on surgical needles. Proc Instn Mech Engrs 211 [H], 335-347. 1997.

(75) Zeltzer D. Autonomy, Interaction and presence. Presence: Teleoperators and Virtual Environments 1, 127-132. 1992.

(76) Newbel D. Guidelines for the Development of Effective and Efficient Procedures for the Assessment of Clinical Competence. Certification and Re-certification of Doctors: Issues in the Assessment of Clinical Competence. Cambridge: Cambridge University Press; 1994. p. 69-91.

(77) Brusilovsky P *et al*. Adaptive Hypertext and Hypermedia. Adaptive Hypertext and Hypermedia. Dordrecht: Kluwer Academic Publishers; 1998.

(78) Southgate L, Jolly B., Bowmer I, Newble D, Norcini J. Determining the content of recertification procedures. In: Newble D, Jolly B., Wakeford R., editors. The Certification and Recertification of Doctors: Issues in the Assessment of Clinical Competence. Cambridge: Cambridge University Press; 1994. p. 178-86.

(79) Van der Vlueten C *et al*. Methods of Assessing and Certification. In: The Certification and Re-certification of Doctors: Issues in the Assessment of Clinical Competence. Newbel DJ, editor. Cambridge: Cambridge University Press; 1994. p.115-25.

(80) Poloniechi J. Half of All Doctors are Below Average. BMJ 316, 1734-1735. 2001.

(81) Poloniecki J, Valencia O, Littlejohns P. Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery. BMJ [316], 1697-1700. 1998.

(82) Klein R. Competence, Professional Self Regulation and the Public Interest. BMJ 316, 1740-1742. 1998.

(83) Campbell MJ. Models, Test and Data. Statistics at square two. London: BMJ Books; p.1-10. 2001.

(84) Castle WM. Fairness in sampling: how to be on target. Statistics in small doses. London: Pearson Professional Limited; p. 85-98. 1995.

(85) Nolte LP, Zamorano L, Visarius H, Berlemann U, Langlotz F, Arm E, *et al*. Clinical Evaluation of a System for Precision Enhancement in Spine Surgery. Clinical Biomechanics 1, 0(6):293-303. 1995.

(86) Grange S. Bunker T, Cooper J. Networking virtual reality for shoulder arthroscopy. British Journal of Healthcare Computing 13(10): p.26-8. 1996.

(87) Lengyel J. The Convergence of Graphics and Vision. Computer 31[7], p.46-53. 1998.

(88) Mayer RE. Multimedia Principle. Multimedia Learning. Cambridge: Cambridge University Press; [1] p.63-80. 2001.

(89) Albanese M. Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. Medical Education; (34):729-38. 2000.

(90) Grange S. The AIPES scoring system for evaluation of surgical simulator performance. p. 64-71. VRIC; France 2004.

(91) Mc Cloy R, Stone R. Science, medicine, and the future: Virtual reality in surgery. BMJ 2001 Oct 20; 323 (7318):912-5.

(92) González MA, Ballester AZ. Combined Statistical and Geometrical 3D Segmentation and Measurement of Brain Structures. Proc. IEEE Workshop on Biomedical Image Analysis; Santa Barbara, Ca: IEEE Computer Society; 1998.

(93) Marias K, *et al*. "Registration and matching of temporal mammograms for detecting abnormalities": Medical Image Understanding and Analysis (MIUA) '99; 1999.

(94) Ota D., Loftin B., Saito T., Lea R., Keller J. Virtual reality in surgical education. Computers in Biology and Medicine 1995; 2 (25):127-37.

(95) Paulo Dario. VOEU Simulation Standards Outline. Personal Communication 2000.

(96) Grange S, Power G, Wills G, Bailey, Miles-Board T, Carr L, Hall W. . Virtual Orthopaedic European University. IAM Group University of Southampton 2002 Available from: URL: http://voeu.ecs.soton.ac.uk

(97) Lowe D. Hypermedia and the Web - An engineering Approach. London: John Wileys & Sons; 1999.

(98) Hall W *et al*. Rethinking Hypermedia The Microcosm Approach .Dordrecht, Netherland, Kluwer Academic Publishers, 1996.

(99) Collins AB *et al*. Cognitive Apprenticeship: Teaching the Craft of Reading, Writing & Mathematics. In: L.B.Resnick, editor. Knowing, Learning and Instruction: Essays in honor of Robert Glaser .Hillsdale, New Jersey: Erlbaum.; 1989.

(100) Nkanginieme KEO. Clinical Diagnosis as a Dynamic Cognitive Process: Application of Bloom's Taxonomy for Educational Objectives in the Cognitive Domain. Med Educ Online 1997; 2(1) Available from: URL: http://www.med-edonline.org/issue2.htm#v2

(101) Research Committee. The Opportunities for Virtual Reality in Simulation. London: The Royal College of Surgeons of England, 2001.

(102) Heilig M, inventor; Sensorama. US patent 050, 870. 1962.

(103) Rubino F**,** Soler L, Marescaux J, Maisonneuve H. Advances in virtual reality are wide ranging BMJ, Mar 2002; 324: 612 ; doi:10.1136/bmj.324.7337.612

(104) Satava RM. Medicine 2001: The King is Dead in: Interactive Technology and the New Paradigm for Health Care. Washington DC: IOS Press; (7) 112-132, 1995.

(105) Satava RM. Virtual reality surgical simulator. The first steps. Surgical Endoscopy 1993; 3 (7):203-5.

(106) Satava RM. Medical applications of virtual reality. Journal of Medical Systems 1 995; 3 (19):275-80. 2001

(107) Satava RM. Emerging medical applications of virtual reality: a surgeon's perspective. Artificial Intelligence in Medicine 1 996; 4 (6) :281-8.

(108) T.Bray *et al*. Extensible Markup Language (XML) 1.0. W3 org; 2, 2000

(109) Organisation. Pub Med. National Institute for Health 2004 Available from: URL: www.pubmedcentral.nih.gov

(110) Dublin Core Metadata Initiative. Dublin Core Committee 2004 Available from: URL: http://dublincore.org/documents/decs

(111) Wactlar H.D. *et al*. Lessons Learned from Building a Terabyte Digital Video Server. Computer 32[2], 66-73. 1999.

(112) S Grange, J Cooper, Bunker T. Exeter Virtual Worlds Shoulder Arthroscopy Simulator. UKVRSIG; London: Brunel Univ Press; 1997.

(113) Hazemi R *et al*. Reinventing the Academy. In: Wilbur S HRHS, editor. The Digital University. London: Springer; p. 7-24. 1998.

(114) Brousseau. *Theory of Didactical Situations in Mathematics*. Dordrecht; 1997.

(115) Krishna *et al*. The myGrid Notification Service. Nottingham, UK 2003 p.475-82.

(116) Grange S, Jones G, Bunker T. Using Java to Embed Complex Simulation Media into Surgical Training Environments. *ISSN 0-7803-6449-X; IEEE EMBS International Conference on Information Technology Applications in Biomedicine*; p.190-6. 2000.

(117) Technical Committee JTC 1/SC34. Hypertext Markup Language. 2000.

(118) David De Roure. Worldwide Universities Network Grid. University of Southampton 2004. Available from: URL: http://wungrid.org

(119) VOEU Consortium. Virtual Orthopaedic European University. IST-13079-1999. European Commission DGXIII 2001.

(120) Holler E. Evolving the ESSENTIAL Project. Personal Communication; Grange S, 1998.

(121) Brett PN. Essential Project proposal. Personal Communication; Grange S, 1998.

(122) David T *et al*. Problem Based Learning in Medicine. 1 London: RSM Press; 1999.

(123) Piaget J. The construction of reality in the child. New York, Basic Books; 1954.

(124) Papert S. Mindstorm. New York: Basic books; 1980.

(125) Wenger E. Communities of practice - learning, meaning and identity. Cambidge: Cambridge University Press; 1998.

(126) Suchman L. Plans and Situated Actions: The Problem of Human/Machine Communication. Cambridge: Cambridge University Press; 1988.

(127) Brown J, Ceal C, Conole G. Situated learning and the culture of learning Educational Researcher; 18(1):32-42. 1989.

(128) Lave EW. Situated Learning: Legitimate Periperal Participation. Cambridge: Cambridge University Press; 1990.

(129) Engestrom Y. Perspectives on activity theory. Learning in doing: social, cognitive and computational perspectives. Cambridge: Cambridge University Press; 1999.

(130) Conole G, Ceal C. A toolkit for supporting evaluation. ALT-J 9[1], 38-49. 2001.

(131) JCHST training committee. Trauma and Orthopaedic Guidance Notes. Joint Committee for Higher Surgical Training 2004 Available from: URL: www.jchst.org/pdf/ortho_ccst_information_sheet.pdf

(132) RCS Fellows. Which of the following are you? The Royal College of Surgeons of England 2004. Available from: URL: www.rcesng.ac.uk/surgical/training

(133) Rossi Mori A, Gangemi A, Steve G. An ontological analysis of surgical deeds. AIME97 p1-11, 1997.

(134) Duchastel P. "Formal and Informal learning with hypermedia". In: D.H.Jonassen and H.Mandl, editor. *Designing Hypermedia for Learning.* Berlin Heidelberg New York: Springer-Verlag; 1989.

(135) G Power, Hall W, Wills G. Realism & Naturalness in a conversational multimodal interface. International Speech Communication Association (ISCA) Tutorial and Research Workshop on Dialogue in Mobile Environments; Germany: Kloster Isree; 2002.

(136) Hart SG, Staveland LE. Development of the NASA TLS: results of empirical and theoretical workload. In: Hancock PA., Meshkayi N., editors. Human Mental Workload. 1 ed. Amsterdam: North-Holland; 1988.

(137) Barfield W., Furness T. Introduction to Virtual Environments. Virtual Environments and Advanced Interface Design. 1st ed. Oxford: Oxford University Press; 1995. p. 3-14.

(138) Grange S., Cooper J., Ponty Y., Bunker T. Comparison of an Electro-optical versus an Electro-magnetic guidance system in the preparation of a Multi-dimensional (6dof) video recorded Minimal Access Surgical environments. University of Exeter; Internal Report No: 386. 1998.

(139) Preece J. What is HCI? Human-Computer interaction. 1[1], 3-27. Harlow, Addison-Wesley. 1998.

(140) Ellison RJ, Fisher DA, Linger RC, Lipson HF, Longstaff TA, Mead NA Survivability: Protecting your critical systems. IEEE Internet Computing 3[6], 55-63. 1999.

(141) SSADM User's Group. SSADM and GUI Design: A Project manager's Guide. 1 ed. London: HMSO; 1994.

(142) Gould J, Dall LC. Designing for usability - key principles and what designers think of them. ACM 28[3], 300-3 11. 1985.

(143) Linda Harasim. A Framework for Online Learning: The Virtual-U. Computer 2004; 32(9):44-9.

(144) Schatz B. CH. Digital Libraries: Technological Advances and Social Impacts. Computer 1999; 82(2):45-50.

(145) IMS Content Information Model. http://www.imsproject.org/content/packaging/cpinfo10.html. 2001.

(146) Gonczi A. Competency Based Assessment in the Professions in Australia. Assessment in Education 1, 27-44. 1994.

(147) Grant J. Review of the effectiveness of CME. In CME - making sure it works. London: www.open.gos.uk/doh/meconf.htm; 1998.

(148) Wills G, Miles-Board T, Bailey C, Carr L, Power G, Hall W, Grange S. The Dynamic Review Journal: a scholarly archive. New Review of Hypermedia and Multimedia 11 [1], 69-89. 2005.

(149) Dumas JS, Redish JC. A Practical Guide to Usability Testing. 1 ed. Portland, OR, Intellect Books; 1999.

(150) Bunker T, Wallace A. Shoulder Arthroscopy. Shoulder Arthroscopy. 1 ed. London: Dunnitz; 1991.

(151) Conole G, *et al*. A pedagogical framework for embedding C and IT into the curriculum. ALT-J 1998; 6 (2):4-16.

(152) William Hersh. The telemedicine curriculum. Journal of Telemedicine and Telecare 2004; 9(6):353-4.

(153) Dublin Core Metadata Initiative. Dublin Core Committee 2004. Available from: URL: http://dublincore.orgdocuments/decs

(154) IMS Content Information Model: IMS Project; 2001. Available at URL: http://www.imsproject.org/content/packaging/cpinfo10.html

(155) Dearing R. The Dearing Report. London: National Committee of Inquiry into Higher Education [NCIHE]; 1997.

(156) University of California Digital Library. UCSD 2004Available from: URL: http://www.cdlib.org/about/faq/

(157) Zhang A., Chang W., Sheikholeslami G., Syeda-Mahmood T. Netview: Integrating Large-Scale Distributed Visual Databases. IEEE Multimedia 5[3], 47-59. 1998.

(158) Jin Jin Zheng, Jian J Zhang. Texture Mapping on Irregular Topology Surface. London: Information Visualisation 02; 2002 p. 323-30.

(159) PACS Systems. GE Medical 2003Available from: URL: http://www.gemedicalsystems.com/caen/it_solutions/rad_pacs/education/pac_ed_index.html

(160) Jolly B *et al*. The reproducibility of assessing radiological reporting: studies from the development of the General Medical Council's Performance Procedures. Med Educ. 35 [supp 1], 36-44. 2001.

(161) Taffinder N, Smith S, Jansen J, Ardehali B, Russell R, Darzi A. Objective measurement of surgical dexterity - validation of the Imperial College Surgical Assessment Device (ICSAD). Minimally Invasive Therapy and Allied Techniques 7 [supp 1], 11. 1998.

(162) Spencer F. Teaching and measuring surgical techniques: the technical evaluation of competence. Bull Amer Coll Surg [63], 9-12. 1978.

(163) Galasko CSB. Assessment of Professional, Clinical and Surgical Skills in the Workplace, Surgical Competence. 1999.

(164) National Electronic Library for Health. More information regarding this resource can be found URL; http://www.nelh.nhs.uk/

(165) Information concerning this project is available from URL: http://www.equator.ecs.soton.ac.uk/projects/

Document identifier:	Simon Grange PhD Thesis
Version number:	Final
Document title:	050406PhD_SGrange
Document type:	PhD Thesis
Local document ref:	C:\Documents and Settings\Simon Grange\Desktop\050406PhD_SGrange.doc
Approval status:	Resubmission Final post Minor Corrections
Keywords:	Education, orthopaedics, pedagogy, digital,
QA information:	Reviewed by Prof Ajit Narayanan, (Professor in Computer Science)
Word Count	81,771
Date of issue:	05/04/2006 14:41
Department:	School of Engineering, Computer Science and Mathematics
Author:	Mr Simon Grange MB ChB FRCS (Tr & Orth)
Telephone:	+44 (0) 784 157 2167
Email:	simon@grange.me.uk
Confidentiality:	University of Exeter, UK.
Distribution:	University of Exeter, UK. +/- British Library
Copyright notice:	University of Exeter, UK.