

Research

The (human) science of medical virtual learning environments

Robert J. Stone*

Human Interface Technologies Team, School of Engineering, University of Birmingham, Birmingham B15 2TT, UK

The uptake of virtual simulation technologies in both military and civilian surgical contexts has been both slow and patchy. The failure of the virtual reality community in the 1990s and early 2000s to deliver affordable and accessible training systems stems not only from an obsessive quest to develop the ‘ultimate’ in so-called ‘immersive’ hardware solutions, from head-mounted displays to large-scale projection theatres, but also from a comprehensive lack of attention to the needs of the end users. While many still perceive the science of simulation to be defined by technological advances, such as computing power, specialized graphics hardware, advanced interactive controllers, displays and so on, the true science underpinning simulation—the science that helps to guarantee the transfer of skills from the simulated to the real—is that of human factors, a well-established discipline that focuses on the abilities and limitations of the end user when designing interactive systems, as opposed to the more commercially explicit components of technology. Based on three surgical simulation case studies, the importance of a human factors approach to the design of appropriate simulation content and interactive hardware for medical simulation is illustrated. The studies demonstrate that it is unnecessary to pursue real-world fidelity in all instances in order to achieve psychological fidelity—the degree to which the simulated tasks reproduce and foster knowledge, skills and behaviours that can be reliably transferred to real-world training applications.

Keywords: human factors; human-centred design; virtual environments; medical and surgical simulation; fidelity

1. INTRODUCTION

The training of future surgeons presents major logistical challenges to healthcare authorities, especially given the constraints imposed by the European Working Time Directive and the consequent reduction in exposure of surgical trainees to real patients and in-theatre experiences. In both the UK and USA, training programmes are now much shorter than ever before [1,2] and, with patients and relatives acquiring detailed knowledge about medical conditions and interventions from the Internet and other highly accessible digital sources, the increasing requirement for consultants to front assessments and interventions is reducing specialized training opportunities for junior medical personnel even further. The training of military surgeons presents an even bigger challenge, especially given the unprecedented levels of trauma they will face [3] and the highly dangerous contexts in which they will have to work. These experiences are highly difficult to train realistically using home territory facilities, although location-based moulage exercises, such as the Ministry of Defence’s hospital exercises (or HOSPEX), contribute significantly to pre-deployment experiences and, to a limited extent, help in the psychological desensitization

process when confronted with severe physical trauma. However, from austere hospital camps in extreme temperatures to forward operating bases, and from medical emergency response teams (MERTs) to the ethical handling non-military casualties of war, there are still significant gaps in the training of surgeons destined for operational military duty.

Many specialists in the world of military medical training believe that there has never been a more pressing need to develop effective simulation-based training to fill these gaps, servicing the needs both of individuals and small teams. However, the form that simulation should take is still a constant source of debate, despite years of experience with simulation throughout both defence and civilian medical communities. Should the simulation be real, based on moulage scenarios with real actors or instrumented mannequins? Or have virtual or synthetic environment technologies today reached a level of maturity and affordability whereby simulation-based systems can guarantee effective knowledge- and/or skills transfer from the point of educational delivery (be that a classroom, the Internet or via hand-held computing platforms) to the real world? Is one form of training more likely to succeed over the other, or should instructors be exploiting both live and virtual contexts? If the latter, how can the balance between the two forms of training be specified?

Regrettably, and despite many decades of high-level experimentation and evaluation of simulators for the

* Author for correspondence.

One contribution of 20 to a Theme Issue ‘Military medicine in the 21st century: pushing the boundaries of combat casualty care’.



Figure 1. Early virtual reality radiography training concept demonstration with virtual lung (University of North Carolina at Chapel Hill).

medical community in general, the answers to these questions are still elusive. To understand why this is so requires a very brief and a critical overview of developments in medical simulation, focusing particularly on the explosion of interactive technologies and virtual environment (VE) or virtual reality (VR) simulators emerging from the early 1990s.

2. MEDICAL SIMULATION AND VIRTUAL REALITY

Space does not permit an in-depth review of developments in medical simulation from the pioneering electromechanical *Sim One* intubation/anaesthesia trainer of the 1960s [4,5] to the large-scale, so-called immersive 'CAVE' multiple display systems being touted by present-day medical simulation centres across the globe, such as the US Uniformed Services University Medical Simulation Center's highly ambitious wide area VE for MERT training [6]. Two papers covering the history of simulation have already been published and are well cited [7,8].

During the 1980s, many visionaries—notably those at the University of North Carolina and within the Department of Defense in the USA—were developing the notion of the *surgeon of the future*. Future medical specialists would, they claimed, be equipped with head-mounted displays (HMDs) and other wearable technologies, rehearsing in VE such procedures as detailed inspections of the unborn foetus or gastrointestinal tract, the accurate targeting of energy in radiation therapy (figure 1), even socket fit testing in total joint replacement. For many years, the US led the field in medical VR. During the late 1980s and early 1990s, interest in the application of VR technologies to medical and surgical training steadily increased, with US-based companies, such as high tech-planations (HT Medical) and Cinémed pursuing the holy grail of 'making surgical simulation real' [9]. Indeed that quest is still evident today. By the mid-1990s, advances in computing technology had certainly developed to support attempts to deliver reasonably interactive anatomical and physiological simulations of the human body (as shown in figure 1). From the digital reconstruction of microtomed convicts (e.g. the Visible Human Project [10], which went on to spawn its own *Visible Human Journal of Endoscopy*), to



Figure 2. Example of Laerdal SimMan mannequin with limb trauma.

speculative deformable models of various organs and vascular systems [11], the quest to deliver comprehensive 'virtual humans' using dynamic visual, tactile, auditory and even olfactory data was relentless. Yet, with a handful of exceptions, the uptake of these simulations by surgical research and teaching organizations was not (and still is not) as widespread as the early proponents of virtual surgery had predicted.

One can partly attribute this apparent failure to a lack of technological appreciation or foresight on the part of individual specialists or administrators within the target medical organizations. However, it was all too easy to forget that most medical organizations outside of the niche simulation centres in the USA and elsewhere simply could not justify the excessive initial costs of the so-called graphics 'supercomputers' and baseline software needed to power these simulations—not to mention crippling annual maintenance charges, depreciation and, in a constantly changing world of information technology, rapid obsolescence.

Little wonder, then, that many simulation centres in today's 'post-VR' era focus their training technology requirements on full- or part-body physical mannequins. Today's instrumented mannequins (e.g. figure 2) are able to fulfil the training needs of medical students from a wide range of backgrounds and specialties—nurses, technicians, paramedics and even pharmacists. Their physiology can be electromechanically controlled in real time, by computer or remote instructor; they can be linked to a variety of typical monitoring devices and are capable of responding to a range of medications, with some systems even delivering symptoms based on the mannequin's 'age' and sex. Many of today's mannequin systems also possess impressive 'after-action review' capabilities, so that events can be replayed and feedback can be provided to trainees on a step-by-step basis, or after a complete scenario has been performed.

But does this mean that digital simulation or VR-like solutions are no longer a player in the world of medical training? Certainly not. Since the VR failures of the closing decade of the twentieth century and early years of the twenty-first century, there have been a number of important developments in the digital simulation arena, the most notable being what

many describe as 'serious games'. Serious games are games 'with a purpose'. In other words, they move beyond entertainment *per se* to deliver engaging interactive media to support learning in its broadest sense [12]. In addition to learning in traditional educational settings, gaming technologies are also being applied to simulation-based training in defence, cultural awareness, in the fields of political and social change and, slowly, to healthcare and specialized medical and surgical training. Another aspect of serious games—accessibility—makes their application potential even stronger than their VR predecessors. Many of today's emerging applications are deliverable via standard archiving media, such as the CD, DVD or memory stick. However, the ever-increasing power of the Web (and Web 2.0 in particular, supporting massively multiplayer online games, or MMOGs) is beginning to offer a highly accessible alternative, both for launching serious games and for recording the performances of individual participants and teams. These qualities are revolutionizing the exploitation of the Web, and will undoubtedly support the execution of unique online learning experiments.

One of the most impressive features of the emerging serious games community is the fact that many of the concept capability demonstrators in existence today are hosted on relatively inexpensive laptops and personal computers (e.g. £500–£700). To achieve today the same simulated visual and behavioural quality that was evident 10–12 years ago would, as hinted earlier, have required a graphics 'supercomputer' costing 20–30 times as much, not to mention software tools costing of the order of £70 000 and above. In addition, tools are becoming available that support the development of interactive three-dimensional (i3D) content by a much wider range of contributors than was the case with VR in the 1990s. Many of these tools—and the software 'engines' that bring their results to life—are made available by mainstream entertainment games companies. While many of the commercially oriented tools and engines demand significant investment for commercially exploitable applications (i.e. applications other than modifications to a game for personal interest or research and development), other, open source systems often permit licence-free development and distribution.

3. SIMULATION AND HUMAN SCIENCE

The issues of cost and accessibility, together with developments in gaming technologies, are strong drivers in today's resurrection of digital simulation exploiting gaming technologies. However, one of the most important, if not *the* most important lesson to be learned from the 1990s is that the failure of the VR community to deliver immediately usable (i.e. non-prototype) digital simulation technologies to the medical world has to be attributed to a poor understanding—more often than not on the part of simulation developers—of the needs, capabilities and limitations of the end user population—the clinical and surgical trainees. Unlike other areas characterized by the adoption of 'high-tech' solutions (such as aerospace, automotive and defence engineering), the input from human factors specialists in the development of

technology-based training systems was not—and is still not today—recognized as an essential precursor to the design of simulation content and the appropriate exploitation of interactive technologies. Human factors is the study (or science) of the interaction between the human and his or her working environment. It makes no difference if that working environment is real or virtual.

One highly relevant study drawing attention to the human factors aspects of i3D simulations—albeit from within the aviation maintenance industry—was conducted by Barnett and others at Boeing's St Louis facility in the late 1990s [13]. Although a number of their conclusions had been anecdotally reported for some time prior to this, their paper added considerable weight to a growing concern in some quarters of the human factors community that the majority of VR and i3D proponents were over-concentrating on technological issues, as opposed to meeting the learning needs of the end user. The same trend is evident today, with early developments in games-based training projects. The project, reported at the 2000 Interservice/Industry Training, Simulation and Education Conference (IITSEC), was performed in support of a virtual maintenance training concept for Boeing's *Joint Strike Fighter* (JSF) proposal. The boeing team set out to evaluate the use of three-dimensional computer-aided design models, implemented within a VE, as a low-cost partial replacement for physical mock-up aircraft maintainer trainers. Participants were required to carry out simple remove-and-replace maintenance procedures using three-dimensional representations of JSF components. One of the results of the early part of this study was that users equipped with so-called 'immersive' VR equipment, namely a HMD and spatially tracked hand controller (or 'wand'), experienced longer training times than those provided with more conventional 'desktop' interactive three-dimensional training methods, and also showed diminished task performance. In particular, Barnett *et al.* [13] noted that '... as a result of... unique features of the VR, four of the participants commented that they focused more on interfacing with the VR than with learning the task...' (p. 154) (Author's emphasis).

The 'unique features' in this case included poor field of view and depth perception offered by the HMD, virtual object distortions and object manipulation artefacts. This is but one example from the VR era and clearly demonstrates that the capability of a simulation-based system to educate and train can be completely destroyed if the simulator content—and simulation *fidelity* in particular—is implemented inappropriately and without a sound human factors underpinning.

4. SIMULATOR FIDELITY MANAGEMENT—THE HUMAN SCIENCE OF SIMULATION

In very general terms, fidelity is a term used to describe the extent to which a simulation represents the real world, including natural or man-made environments and, increasingly, participants or *agents* (including virtual humans or 'avatars'). However, when applied to simulation, it becomes apparent

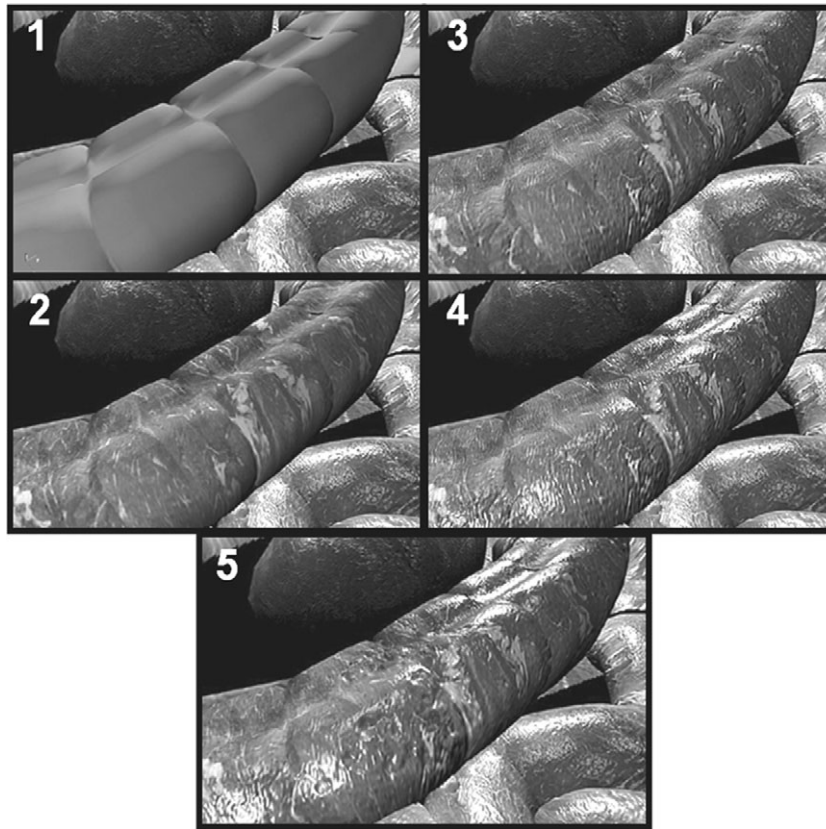


Figure 3. Virtual colon images showing different levels of visual fidelity.

from the literature that there are many variations on the theme of fidelity.

Physical fidelity, or *engineering* fidelity [14], relates to how the VE and its component objects mimic the appearance and the operation of their real-world counterparts. By contrast, *psychological* fidelity can be defined as the degree to which simulated tasks reproduce behaviours that are required for the actual, real-world target application. Psychological fidelity has also been more closely associated with positive transfer of training than physical fidelity and relates to how skills and/or knowledge acquired during the use of the simulation—attention, reaction times, decision making, memory and multi-tasking capabilities—manifest themselves in real-world or real-operational settings.

In far too many examples of simulation design (with particular reference to examples from the medical and surgical community) it has become apparent that physical and psychological fidelities do not necessarily correlate well. More and more physical fidelity does not necessarily guarantee better psychological fidelity. In figure 3, for example, can the same learning and skills transfer (psychological fidelity) be achieved by exploiting the lower physical fidelity virtual human anatomy in this sequence (i.e. images 1 or 2), or those of higher physical fidelity (images 4 or 5, with associated higher costs and longer development times), or via some other representation?

Establishing the components of a task that will ultimately contribute to how psychological fidelity is implemented within a simulation is not an exact science [15]. Observational task analyses need to be conducted with care if those human performance

elements of relevance to defining psychological fidelity are to be isolated effectively. As defined by the present author [16], a human factors task analysis is a process by which one can formally describe the interactions between a human operator and his/her real or virtual working environment (including special-purpose tools or instruments), at a level appropriate to a pre-defined end goal—typically the evaluation of an existing system or the definition of the functional and ergonomic features of a new system. Some examples of the outcome of human factors task analyses in medical and surgical training contexts will be summarized later.

Writing on the subject of simulation for teamwork skills, Beaubien & Baker [17] propose three areas in which fidelity can be manipulated: equipment fidelity, environmental fidelity and psychological fidelity. However, this classification only goes part of the way to define the physical fidelity elements of a simulation that impact upon the psychological fidelity. Experience gained since the mid-1990s in developing VEs and, more recently, games-based simulations for part-task training applications (in a range of defence and medical sectors) suggests that, when observing tasks, there are four key classes of fidelity to consider, each of which impact on defining the ultimate physical and psychological attributes of the simulation [18]. They are:

- *Task fidelity*: the design of appropriate sensory and behavioural features into the end user's task that supports the delivery of the desired learning effect.

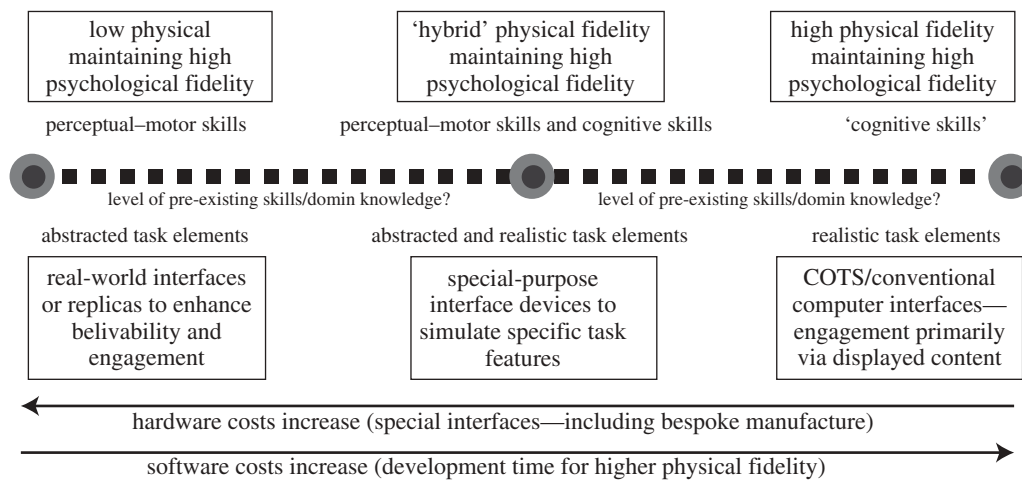


Figure 4. A proposed continuum to support early human-centred decisions relating to simulation design [18].

- *Context fidelity*: the design of appropriate ‘background’ sensory and behavioural detail (including the behaviours of virtual humans or ‘avatars’) to complement—and not interfere with—the task being performed and the learning outcomes.
- *Interactive technology ‘fidelity’*: defined by real-world task observations, interactive technology fidelity is the degree to which input (control) and display technologies need to be representative of real life human–system interfaces. Where they do not, there needs to be careful management of the mapping between the control and display hardware presented to the end user.
- *Hypo- and Hyper-fidelity*: the inclusion of too little, too much or inappropriate sensory and/or behavioural detail (task, context and interaction systems) leading to possible negative effects on serious game/simulation performance and on knowledge or skills transfer.

Building upon this, figure 4 presents a design continuum based on the experience of conducting human factors task analyses in support of applying i3D and serious games technologies to a variety of real-world training domains, and especially to the development of virtual surgical training technologies [18]. In essence, the continuum proposes that, to achieve a successful learning outcome when developing part-task simulators, the design of the simulated tasks and the interaction with those tasks should take into consideration:

- whether the task to be trained is fundamentally *perceptual–motor* (e.g. skills-based) or *cognitive* (e.g. decision-based) in nature (or a combination of the two), and
- whether or not all or only a percentage of the members of a target audience possess pre-existing (task-relevant) perceptual-motor skills and domain knowledge.

The impact of these issues on physical and functional fidelity is of considerable importance, as is their impact on such issues as the choice of hardware

and software and, of course, developmental costs. Human factors task analyses, supplemented with real-world observations and briefings or interviews, should strive to uncover what pre-existing skills and domain knowledge already exist, together with the experience and attitudes of end users to computer-based training technologies.

Turning first to the left-hand extreme of this continuum, if the target training application is primarily perceptual–motor (i.e. skills-based, requiring manual handling, hand–eye coordination or dextrous activities, for instance), or designed to foster simple, sequential procedures (characterized by minimal cognitive effort), then experience shows that endowing the simulator content with high physical fidelity is unnecessary. This statement has to be qualified by stressing that the low(er) fidelity objects and scenes must be designed—again with appropriate human factors input—such that they are reasonable representations of real-world activities, or accurate abstractions of the tasks being trained. They should not, as a result of their visual appearance and/or behaviour, introduce any performance, believability or ‘acceptability’ artefacts, thereby compromising skills transfer from the virtual to the real (or related measures, like skill or knowledge fade). In addition, and as part of the ‘enhancing believability’ process, simulation projects that are characterized by low physical fidelity content will benefit from the exploitation of ‘real-world equivalents’ when it comes to interface devices (i.e. devices that are actually used in the real world, but have been modified for simulator use). There is also an issue here that relates to whether or not the perceptual–motor skills may be better trained using real, as opposed to virtual, physical objects (i.e. part-task medical mannequins). A sound human factors analysis will also help in supporting the choice between real or virtual environments, or a blend of the two.

5. CASE STUDY 1—THE MINIMALLY INVASIVE SURGICAL TRAINER

A good medical example of the low-physical, high-psychological fidelity debate is the original minimally invasive surgical trainer (MIST), which evolved from a comprehensive in-theatre human factors task analysis

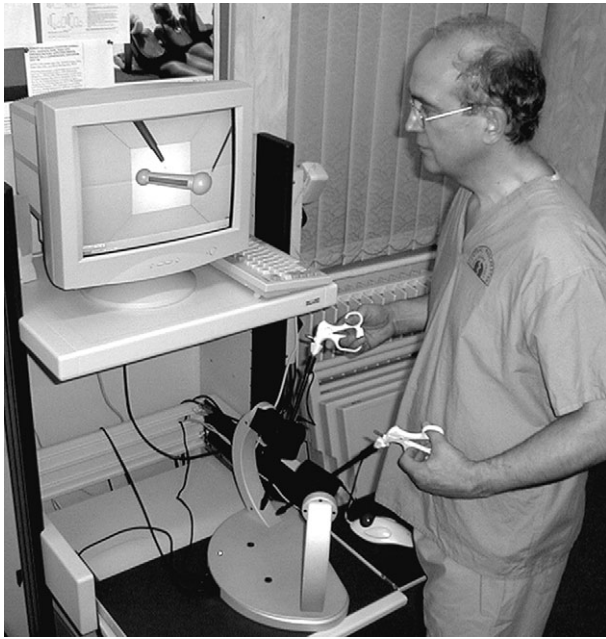


Figure 5. Original minimally invasive surgical trainer (MIST) and laparoscopic gynaecology task module.

in the early 1990s [19,20]. MIST (today marketed by Mentice of Sweden) is a personal computer-based ‘keyhole’, surgical skills trainer for laparoscopic cholecystectomy and gynaecology [21,22]. The original MIST system (figure 5) presented the trainees not with high-fidelity three-dimensional human anatomy and physiology but with visually and functionally simplistic objects—graphical spheres, cubes, cylinders and wireframe volumes—all abstracted from an observational task analysis of surgical procedures evident in theatre (e.g. clamping, diathermy, tissue sectioning, etc.). The analysis made it possible to isolate eight key task sequences common to a wide range of laparoscopic cholecystectomy and gynaecological interventions and then to define how those sequences might be modified or constrained by such factors as the type of instrument used, the need for object or tissue transfer between instruments, the need for extra surgical assistance and so on. By adopting an abstracted-task design process, the MIST simulation content was developed by the present author to avoid the potentially distracting effects (and negative skills transfer) evident with poorly implemented virtual humans at the time [12]. For well over a decade, and with more clinical and experimental evaluation studies than any other simulation-based surgical skills trainer, MIST helped to train the perceptual-motor skills necessary to conduct basic laparoscopic manoeuvres involved in cholecystectomy and gynaecological minimally invasive interventions. ‘Believability’ was enhanced through the use of realistic (and instrumented) laparoscopic instruments which, together with offline video sequences of actual operations, helped to relate the abstracted task elements to real surgical interventions. In many respects, the MIST trainer set the standard for a range of subsequent basic surgical skills trainers (including Surgical Science’s *LapSim* and Symbionix’s *LAP Mentor* products).

Turning now to the right-hand extreme of the continuum in figure 4, if the target training application is primarily cognitive (e.g. requiring timely decisions based on multiple sensory inputs, reaching threshold levels of knowledge to allow progress on to the next stage of a complex process, strong spatial awareness, effective team coordination, etc.), then experience shows that it is important to endow the simulator content with appropriate, but typically *high* physical fidelity in order to preserve high psychological fidelity. However, because the end user is assumed to possess the basic perceptual-motor and manual handling skills that underpin the decision-based training *already* (or if not, then they are being fostered elsewhere, possibly exploiting physical training facilities), then any manual or dextrous activities features in the simulator should be committed to decision-triggered animation sequences. Consequently, the need for realistic (real-world equivalent) interface devices is typically minimized (although it is accepted that there will be exceptions). This, of course, reduces the potentially very high cost of having to modify surgical instruments such that their usage interfaces seamlessly with the simulation. It also means that simulation designers are less tempted to interface the simulation content with outlandish off-the-shelf control products, such as multi-function joysticks, VR gloves and the like (although, again, there have been exceptions, such as the highly publicized use, in 2008 and 2009 of the Nintendo *Wii* controller to ‘foster surgical skills’ [23]).

6. CASE STUDY 2—THE INTERACTIVE TRAUMA TRAINER

A good example of this is the interactive trauma trainer (ITT) [12]. The ITT project came about following requests from defence surgical teams who expressed a desire to exploit low-cost, part-task simulations of surgical procedures for combat casualty care, especially for refresher or *just-in-time* training of non-trauma surgeons who might be facing front-line operations for the first time. The ITT was also the result of an intensive human factors project based on observational analyses and briefings conducted with Royal Centre for Defence Medicine and Army Field Hospital specialists. These analyses contributed not only to the definition of learning outcomes and performance metrics but also to key design features of the simulation and human-computer interface. The task of the user was to make appropriate decisions relating to the urgent treatment of an incoming virtual casualty with a ‘zone 1’ neck fragmentation wound. Appropriate interventions—oxygen provision, blood sampling, ‘hands-on’ body checks, patient visual and physiological observation, endotracheal intubation, and so on—had to be applied in less than 5 minutes in order to save the virtual casualty’s life. However, as stressed earlier, rather than replicate the dextrous surgical handling skills the user would already possess, the simulator would enhance the *decision-making* skills on the part of the surgeon. Consequently, the ITT not only exploited powerful games engine software technology,



Figure 6. The virtual military casualty in the interactive trauma trainer (ITT) concept demonstrator.

supporting the rendering of high-fidelity models of the virtual casualty, it also exploited a typically simple gaming interface—mouse control for view-point change, option selection and instrument acquisition. The human factors analysis helped to define the end shape and form of the ITT, applying high fidelity effects only where they would add value to the surgeon's task (figure 6). The analysis also ensured that the dextrous tasks (e.g. the use of a laryngoscope, stethoscope, intubation tubes, Foley's catheter, etc.) were committed to clear and meaningful *animation* sequences, rather than expecting the surgical users to interact with three-dimensional models of instruments via inappropriate control products with limited or no haptic feedback.

The ITT project also provided some excellent examples of 'hyper-fidelity', as defined earlier). Working in collaboration with a well-known British computer games company, it became evident during the later reviews of the simulation that features of the virtual task and context had been included, not because they were central to the delivery of the learning requirements, but because they were features the simulation developers believed would endow the system with impressive levels of realism. Two examples in particular attracted (or, rather, distracted) the attention of surgical end users during the execution of the basic patient checks and intubation procedures [18]. In the first instance, the simulation developers had programmed exaggerated flexing behaviours into the tubing of the virtual stethoscope. During the airway-breathing-circulation (ABC) animation sequence, and as the animated surgeon's hand moved the stethoscope towards the casualty, the tube flailed dramatically around the scene, which many end users found both comical and distracting. The second example related to the three-dimensional model of the laryngoscope, where the virtual reflective qualities of the surface of the instrument were so intense, they were actually distracting for most users. Furthermore, some users also noted—having stopped the intubation procedure—that the reflection mapped onto the laryngoscope surface was actually wrong showing the casualty's body and scrub nurse—both located on the opposite side of the laryngoscope—instead of the surgeon's own virtual body. During

early evaluations, it was found that almost 90 s of distraction were caused by hyper-fidelity (i.e. 30% of the simulated scenario duration).

The lessons learned during the execution of the ITT project were also exploited in an advanced interactive simulation system for defence medics called *Pulse!!* [24]. Coordinated by Texas A&M University Corpus Christi. Funded in 2006 by a \$4.3 million federal grant from the Department of the Navy's Office of Naval Research, the *Pulse!!* 'Virtual Learning Space' healthcare initiative is designed to provide an interactive, VE in which civilian and military health care professionals can practise clinical skills in order to better respond to catastrophic incidents, such as bioterrorism. The *Pulse!!* training system is based on virtual scenarios as diverse as a realistic representation of Bethesda Naval Hospital in Washington, DC to a busy clinic populated with patients suffering from exposure to anthrax. Medical trainees interact with virtual patients and other medical personnel to conduct examinations, order tests and administer medication.

7. CASE STUDY 3—THE IERAPSI TEMPORAL BONE INTERVENTION SIMULATOR

Finally and turning to the central, 'hybrid' physical fidelity portion of the continuum in figure 4, this situation refers to instances where a task analysis highlights the need for a simulator to possess higher physical fidelity in one sensory attribute over another. In such a case, it may become necessary to develop, procure and/or modify special-purpose interfaces in order to ensure the stimuli defined by the analysis as being essential in the development of skills or knowledge are presented to the end user using appropriate technologies. Take, for example, a medical (mastoidectomy/temporal bone intervention) simulator developed as part of a European Union-funded project called *Integrated Environment for Rehearsal and Planning of Surgical Interventions* (IERAPSI) [25]. Here, the task analysis undertaken while observing ear, nose and throat (ENT) surgeons, together with actual 'hands-on' experience using a cadaveric temporal bone, demonstrated that the skills to be trained were mostly perceptual-motor in nature, but the decisions were also complex (and safety-critical), albeit at limited stages of the task. This drove the decision to adopt a hybrid physical fidelity solution (figure 7) to train mastoid drilling and burring skills based on:

- a low-physical fidelity visual representation of the temporal bone region (omitting any features relating to the remaining skull areas, middle/inner ear structures or other structures, such as the sigmoid sinus and facial nerve);
- a high-fidelity software simulation reproducing the physical and volumetric effects of penetrating different layers of hard mastoid cortex and air-filled petrous bone with a high-speed drill;
- an interface consisting of a binocular viewing system and two commercial off-the-shelf haptic feedback stylus-like hand controllers, capable of

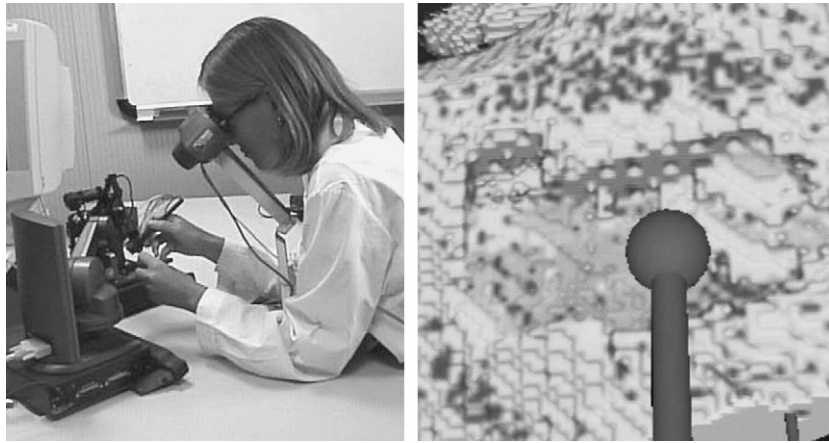


Figure 7. Stereoscopic display and haptic feedback interfaces for mastoidectomy simulator. Image also shows burring activities with a virtual temporal bone model.

reproducing the force and tactile sensations associated with mastoidectomy and the vibration-induced sound effects experienced when drilling through different densities of bone.

A MIST-like low-physical fidelity solution was also investigated, based on a multi-layer volumetric abstraction of the bone penetration process using drills and other medical tool representations. However, concerns were raised regarding the low psychological fidelity inherent in a solution that ignored the significance of stereoscopic vision, haptic feedback and sound effects in the training process. All three of these effects are crucial to the safe execution of a mastoid drilling procedure, helping the surgeon to avoid drill over-penetration and the inadvertent destruction of key nerve paths and blood vessels.

From a cost perspective, real-world or replica interface devices, as used with MIST, are relatively easy to integrate with simplified three-dimensional task representations, although initial bespoke development and manufacturing costs can be high. However, if special-purpose interfaces are to be exploited, then the cost of programming will inevitably increase as attempts are made to generate believable sensory effects (visual, sound, haptics, etc.) that feature in both the simulation content and in the device drivers of the interface equipment.

8. CONCLUSIONS

Although they were reviewing the application of VE techniques in mental health and rehabilitation, the comments of Manchester University's Gregg & Tarrier are just as applicable to medical and surgical simulation, namely that the effectiveness of simulation-based therapy over traditional therapeutic approaches is hampered by a lack of good-quality research [26]. Gregg & Tarrier go on to state that 'before clinicians will be able to make effective use of this emerging technology, greater emphasis must be placed on controlled trials with clinically identified populations' (p. 343). Simulation design *must* become a science, a *human-centred* science, driven by the publication of good,

peer-reviewed case studies with experimental evidence and underpinned by the development of meaningful human-centred metrics for the measurement of fidelity, such as eye tracking, pupillometry, EEG and ECG correlates of performance, skin conductivity, cortisol levels and so on. At the time of writing, innovative approaches to developing integrated psychophysiological sensor suites are being developed at the author's own institution and are referenced elsewhere in the human-computer science literature [27]. Also at the present time, and for the foreseeable future, the design and management of fidelity for digital simulators will remain an inexact science, although case studies are now emerging that will, no doubt, contribute to the growing body of knowledge relating to simulation fidelity.

Another key human factors issue in simulation design is that of 'managing expectations'. One of the major challenges faced by designers of simulators is that they often find themselves having to educate the end users (or system procurers) to appreciate and understand the link between training needs and simulation design, especially if those end users hold up mainstream videogames or military and civilian flight simulator technologies as their 'baseline' concept for excellence. This can lead to a 'fidelity expectation gap' [28] where reality becomes the 'fidelity standard' and the financial costs of the simulation, not to mention the potentially adverse impacts on human performance, increase dramatically. As was witnessed in the ITT project described earlier, the issue here is that if the definition of reality is left to the imagination of the developers of the simulation—the three-dimensional modellers, artists and software coders—or the hyper-reality demands of the culture in which they work (such as the mainstream gaming community), then, from a training perspective, this becomes a potential recipe for failure. Without a clear human-centred approach to the design of future digital simulators for the medical domain, the features that should promote effective learning and knowledge or skills transfer to the real world [29] run a very high risk of being suppressed, even omitted altogether.

As mentioned earlier, the discipline of human factors has had a long and a distinguished history of

influencing the uptake of human-centred design processes in different application domains, such as the automotive industry and defence, and, to a lesser extent (regrettably) medicine and surgery. With ever-increasing complexity in today's computer-based systems, ergonomics can also be looked upon as a 'bridge' between human behaviour and technology, striving to guarantee the usability of future devices and computer-based systems. It follows, then, that it should make no difference to human factors specialists if the working environments they are dealing with are real or virtual. The issues are the same. As stated—over 30 years ago—by Ramon Berguer, Professor of Surgery at the University of California-Davis and one of the pioneers in the application of ergonomics principles to surgical practice:

'A scientific and ergonomic approach to the analysis of the operating room environment and the performance and workload characteristics of members of the modern surgical team can provide a rational basis for maximizing the efficiency and safety of our increasingly technology-dependent surgical procedures' [30, p. 1011].

One issue is clear. Digital simulation, VE, serious games—whatever the field may be called—will never totally replace the need for clinical and surgical training exploiting physical mannequins (which are increasing in sophistication on an almost monthly basis) or live moulage exercises with actors. The technology certainly has the potential to supplement physical and live training as part of a blended 'live-virtual' solution, helping trainees to prepare and rehearse using virtual contexts, but will never replace the live experience altogether. The success of the technology in establishing a significant role in the future of general and specialist medical training depends on a clear understanding that a reliance on technology *per se* will invariably result in failure to deliver a solution that is both future-proof and, most importantly, fit for human use.

REFERENCES

- Morris-Stiff, G. J., Sarasin, S., Edwards, P., Lewis, W. G. & Lewis, M. H. 2005 The European working time directive: one for all and all for one? *Surgery* **137**, 293–297.
- Whalen, T. V. 2007 Duty hours restrictions: how will this affect the surgeon of the future? *Am. Surg.* **73**, 140–142.
- Nessen, S. C., Lounsbury, D. E. & Hetz, S. P. (eds) 2008 *War surgery in Afghanistan and Iraq: a series of cases, 2003–2007*. Textbooks of Military Medicine; Department of the Army, Fort Detrick, MD: Walter Reed Army Medical Center Borden Institute.
- Denson, J. S. & Abrahamson, S. 1969 A computer-controlled patient simulator. *J. Am. Med. Assoc.* **208**, 504–508. (doi:10.1001/jama.208.3.504)
- Gaba, D. 2004 The future of simulation in healthcare. *Qual. Saf. Health Care* **13**, 2–10.
- Ha Lee, C., Liu, A., del Castillo, S., Bowyer, M., Alverson, D., Muniz, G. & Caudell, T. P. 2007 Towards an immersive virtual environment for medical team training. In *Medicine meets virtual reality 15: in vivo, in vitro, in silico: designing the next in medicine*, vol. 125 (eds J. D. Westwood *et al.*), pp. 274–279. Amsterdam, The Netherlands: Studies in Health Technology and Informatics. The IOS Press.
- Bradley, P. 2005 The history of simulation in medical education and possible future directions. *Med. Educ.* **40**, 254–262. (doi:10.1111/j.1365-2929.2006.02394.x)
- Rosen, K. 2000 The history of medical simulation. *J. Crit. Care* **23**, 157–166. (doi:10.1016/j.jcrit.2007.12.004)
- Meglan, D. 1996 Making surgical simulation real. *Comput. Graph.* **30**, 37–39. (doi:10.1145/240806.240811)
- Banvard, R. A. 2002 The visible human project image data set from inception to completion and beyond. In *Proc. of CODATA 2002: Frontiers of Scientific and Technical Data*. Canada: Track I-D-2: Medical and Health Data; Montreal.
- Delingette, H. 1998 Initialization of deformable models from 3D data. In *Proc. of the Sixth Int. Conf. on Computer Vision (ICCV'98)*, pp. 311–316. Bombay, India. Washington, DC: IEEE Computer Society.
- Stone, R. J. & Barker, P. 2006 Serious gaming: a new generation of virtual simulation technologies for defence medicine and surgery. *Int. Rev. Armed Forces Med. Serv.* **79**, 120–128. (June Edition).
- Barnett, B., Helbing, K., Hancock, G., Heininger, R. & Perrin, B. 2000. An evaluation of the training effectiveness of virtual environments. In *Proc. of the Interservice/Industry Training, Simulation and Education Conf., I/ITSEC, Orlando, Florida, 27–30 November 2000*. St. Petersburg, FL: Simulation Systems and Applications, Inc. (Information Engineering Division).
- Miller, R. B. 1954 *Psychological considerations in the designs of training equipment*. Columbus, OH: Wright Air Development Center Publication, Wright Patterson Air Force Base.
- Tsang, P. S. & Vidulich, M. 2003 *Principles and practice of aviation psychology*. Human factors in transportation series. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stone, R. J. 2001 Virtual reality in the real world: a personal reflection on 12 years of human-centred endeavour. In *Proc. of the Int. Conf. on Artificial Reality and Tele-existence*, pp. 23–32. Tokyo, Japan: ICAT.
- Beaubien, J. M. & Baker, D. P. 2004 The use of simulation for training teamwork skills in health care: How low can you go? *Qual. Saf. Health Care* **13**(Suppl. 1), 151–156.
- Stone, R. J. 2008 *Human factors guidelines for interactive 3D and games-based training systems design: edition 1*. Human Factors Integration Defence Technology Centre Publication. See www.hfidtc.com.
- Stone, R. J. 1999 The opportunities for virtual reality and simulation in the training and assessment of technical surgical skills. In *Proc. of Surgical Competence: Challenges of Assessment in Training and Practice*, pp. 109–125. London, UK: Royal College of Surgeons and Smith & Nephew Conference.
- Stone, R. J. & McCloy, R. F. 2004 Ergonomics in medicine and surgery. *Bri. Med. J.* **328**, 1115–1118. (doi:10.1136/bmj.328.7448.1115)
- Gallagher, A. G., McClure, N., McGuigan, J., Crothers, I. & Browning, J. (eds) 1998 Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer, virtual reality (MIST_{VR}). In *Medicine meets virtual reality 6*, vol. 50 (eds J. D. Westwood *et al.*). Amsterdam, The Netherlands: IOS Press, Studies in Health Technology and Informatics.
- Taffinder, N., McManus, I., Jansen, J., Russell, R. & Darzi, A. 1998 An Objective assessment of surgeons' psychomotor skills: validation of the MIST_{VR} laparoscopic simulator. *Bri. J. Surg.* **85**(Suppl. 1), 75.
- Reilly, M. 2008 A Wii warm-up hones surgical skills. *New Sci.* **2639**, 24.

- 24 Johnston, C. L. & Whatley, D. (eds) 2005 Pulse!!—a virtual learning space project. In *Medicine meets virtual reality 14: accelerating change in healthcare: next medical toolkit*, vol. 119 (eds J. D. Westwood, R. S. Haluck, H. M. Hoffman *et al.*), pp. 240–242. Amsterdam, The Netherlands: Studies in Health Technology and Informatics. IOS Press.
- 25 John, N. W. *et al.* 2000 Unpublished IERAPSI report: surgical procedures and implementation specification (Deliverable D2). EU Project IERAPSI (IST-1999-12175). Manchester, UK: University of Manchester.
- 26 Gregg, L. & Tarrier, N. 2007 Virtual reality in mental health: a review of the literature. *Soc. Psychiatry Psychiatric Epidemiol.* **42**, 343–354. (doi:10.1007/s00127-007-0173-4)
- 27 Vice, J. M., Lathan, C., Lockerd, A. D. & Hitt, J. M. 2007 Simulation fidelity design informed by physiologically-based measurement tools. In *Foundations of augmented cognition*. Lecture Notes in Computer Science, no. 4565, pp. 186–194. Berlin, Germany: Springer.
- 28 Northam, G. 2000. Simulation fidelity—getting in touch with reality. In *Proc. of SimTecT 2000, Melbourne, February 2000*. Lindfield, Australia: Simulation Industry Association of Australia (SIAA).
- 29 Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., Gordon, D. L. & Scalese, R. J. 2005 Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med. Teach.* **27**, 10–28.
- 30 Berguer, R. 1999 Surgery and ergonomics. *Arch. Surg.* **134**, 1011–1016. (doi:10.1001/archsurg.134.9.1011)