16

Virtual Reality Simulators for Objective Evaluation on Laparoscopic Surgery: Current Trends and Benefits.

Ignacio Oropesa[1,2], Pablo Lamata[1], Patricia Sánchez-González[1,2], José B. Pagador[3], María E. García[1], Francisco M. Sánchez-Margallo[3] and Enrique J. Gómez[1,2]
[1] Bioengineering and Telemedicine Group, (ETSIT, Technical University of Madrid),
[2]Networking Research Center on Bioengineering, Biomaterials and Nanomedicine,
[3]Minimally Invasive Surgery Centre Jesús Usón,
Spain

1. Introduction

1.1 Laparoscopic surgery

Minimally Invasive Surgery (MIS) has changed the way surgery is performed in Operating Rooms (OR). MIS techniques are increasing their relevance in almost all surgical specialities, and have become the recommended standard in many procedures, displacing open surgery. Laparoscopy, one of the most common MIS approaches, has been adopted by several surgical sub-specialties including gastrointestinal, gynaecological and urological surgery (Fig. 1). It has become the standard technique for certain pathologies, like those associated with anti-reflux diseases, and procedures such as cholecystectomy (Cuschieri, 2005).

Fig. 1. Operating theatre view during a laparoscopic surgical intervention

MIS techniques bring important advantages to patients, such as less postoperative complications, faster recoveries or shorter hospitalization periods. However, they also bring forth considerable limitations and changes for physicians. Specifically, need of high degree of manual dexterity, complexity of instrument control, difficult hand-eye co-ordination, and lack of tactile perception, are the major obstacles. These difficulties involve a challenge for surgeons in getting used to a reduced workspace and to a limited sensory interaction, caused by indirect manipulation and visualization of the patient. They have to acquire new cognitive and motor skills, and they have to accommodate to the reduced workspace and to visualizing the intervention through a 2D monitor.

Due to these limitations, acquisition of MIS skills requires a long learning curve. Moreover, there is also a crescent pressure for safer, transparent and reproducible training programs. They should also allow for practice anywhere at any time, and make use of structured and objective training curricula to determine accurately the trainee´s preparation.

1.2 A historical framework on surgical evaluation

Effective training and assessment of surgeons have become one of the major concerns of hospitals and clinics in recent decades, fuelled mostly by patients' and society's demand for safer surgeries and prepared professionals. Much focus is thus set on the goal of developing structured curricula for surgical qualification and excellence.

1.2.1 Theoretical background

In order to understand the implications of designing training and assessment curricula, it is necessary to bear in mind some of the different pedagogical models and theories for adult learning. More specifically, we will consider the Bloom taxonomy (Bloom et al., 1956) and the Miller pyramid (Miller, 1990).

According to the learning objectives of a training program, Bloom's taxonomy defines three categories of learning objectives: knowledge, skills and attitudes. *Knowledge* refers to cognitive aspects, the assimilation and transformation of information; *skills* to psychomotor competences; and *attitudes* to the growth in feelings or emotional areas.

Most important in clinical education, however, is Miller's pyramid, which establishes four training levels: (1) *Knows* (knowledge), (2) *Knows How* (competence), (3) *Shows How* (performance) and (4) *Does* (action). The first two levels deal with declarative knowledge (knowing *what* to do), and thus can be established by means of examinations or essays. The two top levels are related to procedural knowledge (knowing *how* to do it), where establishment of proficiency levels is not so obvious due to the complex mixture of cognitive, motor, judgment and emotional skills involved.

In a broad sense, these models and theories convey the existence of a double plane of skills to be acquired: cognitive and motor skills. A third level could be arguably considered, involving the trainee's own judgement and applied knowledge to the problem at hand. Whilst in surgery cognitive skills' evaluation can easily be determined by validated methods such as examinations, motor and judgement skills are not so easily established. Thus they have been the focus of attention on recent years, implying the need for standardised and objective training programs (Tavakol et al., 2008).

1.2.2 Towards objectively structured curricula

Traditionally, training of surgeons has been based on the mentor-trainee relationship known as Halsted's model (Halsted, 1904). Motor skills' evaluation is performed with periodic In-Training Evaluation Reports (ITERs), along with aspects such as patient care, communication skills or professionalism (Sidhu et al., 2004). However, these reports are subjective, expensive, and prone to two undesirable side effects: The first is the halo effect, which refers to the influence that the relationship with a trainee can have on the mentor's report, for good or bad. Secondly, as these reports are periodically written, they are subject to the evaluators' long term memory, and provide little or none constructive feedback to the trainee (Fried & Feldman, 2008).

A need for structured, objective curricula was thus detected, and one of the first efforts to remedy the situation were the Objective Structured Clinical Examinations (OSCE), introduced on 1975 by Harden et al., and developed together between the University of Dundee and the Western Infirmary of Glasgow (Scotland). OSCE established a report based on trainee's performance on different clinical stations by means of checklists and assessment reports, with the process and end-product analysis of the task clearly distinguished (Harden et al., 1975).

As successful as OSCEs were, technical skill evaluation is buried between its much more ambitious examination goals, focused on other aspects such as procedural knowledge or attitude towards the patient. As a result of this, and in the wake of their popularity, the Objective Structured Assessment of Technical Skills (OSATS) were developed (Martin et al., 1997). Like OSCE, they employ assessment techniques such as operation checklists and end product-analysis, but always centred on the technical and motor skills of the surgeon. They are usually employed in laboratory settings, using box trainers or human cadavers, and ultimately, live animals in the OR (Sidhu et al., 2004). A counterpart of OSATS for Minimally Invasive Surgery was developed by Vassiliou et al., the Global Operative Assessment of Laparoscopic Skills (GOALS) (Vassiliou et al., 2005).

OSATS validity has been fully established, for skill training ranging from simple tasks to advanced chores (Moorthy et al., 2003). However, the resources needed are high, ranging from the number of experts required at each station to evaluate the trainees, to the marginal costs of each exam per candidate. Laparoscopic video offline-evaluation has been proposed to reduce some of these costs (Datta et al., 2002) with good reliability results; but still the presence of a reviewer is required, and immediate feedback is lost for the trainee.

On the last few years, there has been a been a growing interest on researching automatic methods for measuring the surgeon's motor skills; to provide him precise and immediate feedback on his performance, without requiring the constant presence of a supervisor. Training methods are being gradually changed, leaving the traditional ways behind on behalf of criterion-based curricula (Satava, 2008). This tendency has been boosted thanks to the development and advances on tracking and computing technologies, which have lead, for example, to the appearance of Virtual Reality simulators for surgical training. A new vast research field has opened, were efforts focus not only on the development of training and assessment systems such as said simulators; but on determining what these systems should measure and how should that information be handled (Lamata, 2006a). The present chapter will present an in-depth view on how Virtual Reality simulators have steadily become a part of motor skill formation programs on Minimally Invasive Surgery.

2. Metrics definition: How is surgical skill defined?

2.1 Difficulties on metrics definition

Since the need for structured and objective assessment programs became apparent, there has been much research to identify which parameters characterise surgical skills. This research has been boosted with the availability of automated tracking and registering systems such as simulators. Number, precision and accuracy of the potential metrics have grown due to the processing capabilities these systems provide (Satava et al., 2003).

Determination of valid metrics is a complex process where great difficulties arise when defining quantitatively a surgical motor skill, considering how relative that definition can

be. Different validation studies for a given metric may vary on the conclusions obtained; often, due to the nature and difficulty of the task associated to it. Definition must thus be carefully considered, as some metrics also require *ad hoc* characterization for a given task. Error-related metrics, for example, will be closely associated to that task's goal.
Metrics must be taken into account in relationship with each other rather than on their own. For example, time taken on a task is not a valid parameter if the trainee commits many errors during the exercise. When considering several metrics, it has to be regarded that their nature may vary, and thus also the means for registering them. One may consider, for example, a movement tracking device to capture the tool's motion for analysis of the path length; but to combine that information with input on the errors performed, it will be necessary a supervisor or a post-exercise video-review. In this sense, Virtual Reality simulators excel themselves, due to their ability to determine qualitative parameters such as errors committed and perform final-product analysis.

2.2 Metrics taxonomy

Much research has been devoted to the definition of new valid metrics for performance assessment (Cotin et al., 2002; Lamata, 2006a), as well as on determining the ideal skills, tasks and parameters to measure (Satava et al., 2003). Metrics can generally be classified into two main categories: Efficiency and Quality metrics (Fried & Feldman, 2008). Some of the most important metrics identified in the literature are shown on Table 1.

Metrics for objective skills' assessment			
Efficiency	*Force Analysis*	Tool – tissue forces Torsion	(Rosen et al., 2002)
		Force sensitivity	(Lamata, 2006a)
	Motion Analysis	Time Path Length Motion Smoothness Depth Perception Tool Rotation	(Cotin et al., 2002)
		Speed & Acceleration Optimal Path Deviation IAV (Energy Expenditure)	(Cavallo et al., 2005)
		Angular Area Volume	(Chmarra et al., 2010)
Quality		Task Outcome Errors Manoeuvres' Repetitions Manoeuvres' Order Idle States	(Fried & Feldman, 2008)

Table 1. Metrics for objective skills' assessment

For meaningful skill assessment, both efficiency and qualitative metrics should always be considered on any training curricula. Efficiency metrics are related with measurable physical parameters, and thus their definition is usually precise and has a strong theoretical background behind them. These metrics always require the use of some sensor-based device in order to be acquired, either on physical or virtual simulators; and thus are objective, reproducible and little prone to misinterpretation. A distinction can be made between motion- and force- derived metrics. The first ones include all those related with movements of hands and tools performed during a task: total path length, economy of movements, speed, motion smoothness, etc. Force related metrics, such as tool-tissue interactions, have also been studied by Rosen et al. (Rosen et al., 2002), and, more recently, by Horeman et al. (Horeman et al., 2010).

Quality metrics, on the other hand, relate to the task's definition and execution. Most prominent among these metrics are the errors committed, the final product analysis, the sequence of steps performed in an exercise or procedure, etc. Objective and automatic measurements of these parameters can be difficult, and usually they call for the presence of a trained supervisor and the definition of clear structured checklists, such as those provided by OSATS (Fried & Feldman, 2008).

2.3 Validation of metrics for skills' assessment

As shown above, there are many potential metrics to be considered for surgical assessment; however, not all of them prove to be as decisive for the task. A process of validation must be carried out in order to determine their relevance and suitability for the evaluation process.

2.3.1 Concepts on validation

In order for a test, or measurement within it, to be considered useful for the determination of surgical skills, proof of its reliability and validity must be given (Fried & Feldman, 2008).

Reliability is a measure of the consistency of the test; the extent to which the assessment tool delivers the same results when used repeatedly under similar conditions. It is measured by a reliability coefficient, quantitative expression of the consistency of the tests ranging between 0 and 1. A good reliability coefficient has been approximated at values >0.8. Other useful measures of reliability are α, coefficient α, Cronbach's α, or internal consistency (Gallagher et al., 2003). Three different aspects are involved:

- **Inter-rater Reliability:** Extent to which two different evaluators give the same score in a test made by a user. This feature has little interest in Virtual Reality simulators, where metrics are already automatically acquired.
- **Intra-rater Reliability:** Internal consistency of an evaluator when grading on a given test on different occasions.
- **Test-retest Reliability:** Extent to which two different tests made by the same person in two different time frames give the same result.

Validity relates to the property of "being true, correct, and in conformity with reality". In testing, the fundamental property of any measuring instrument, device, or test is that it ''measures what it purports to measure''. Within the testing literature, a number of benchmarks have been developed to assess the validity of a test or testing instrument. They are the following (Gallagher et al., 2003):

- **Face validity:** defined as ''a type of validity that is assessed by having experts review the contents of a test to see if it seems appropriate''. It is a very subjective type of validation and is usually used only during the initial phases of test construction. For example a simulator has face validity when the chosen tasks resemble those that are performed during a surgical task.

- **Content validity:** defined as ''an estimate of the validity of a testing instrument based on a detailed examination of the contents of the test items''. Experts perform a thorough examination of the contents of the tests to determine if they are appropriate and situation specific. Establishing content validity is also a largely subjective operation and relies on the judgments of experts about the relevance of the materials used. For example a simulator has content validity when the tasks for measuring psychomotor skills are actually measuring those skills and not anatomic knowledge.
- **Construct validity:** degree to which the test captures the hypothetical quality it was designed to measure. A common example is the ability of an assessment tool to differentiate between experts and novices performing a given task (Schijven & Jakimowicz, 2003).
- **Concurrent validity:** defined as ''the extent to which the test scores and the scores on another instrument purposing to measure the same construct are related''. When the other instrument is considered a standard or criterion, the validity test is called "criterion validity" Discriminate validity is defined as ''an evaluation that reflects the extent to which the scores generated by the assessment tool actually correlate with factors with which they should correlate''.
- **Predictive validity:** defined as ''the extent to which the scores on a test are predictive of actual performance''. An assessment tool used to measure surgical skills will have predictive validity if it can ascertain who will perform surgical tasks well and who will not.

2.3.2 State of the art

Despite all of the metrics stated previously, many of them still require proper validation in order to be considered representative of surgical skill level. Time, total path length and economy of movements are considered in general as valid metrics, on the basis that an expert surgeon will perform a task more swiftly and denoting a more clear perception of the surgical space and the strategic approach to the task at hand (Thijssen & Schijven, 2010). Path deviation is also a very popular metric, usually considering the optimal path as the straight line between two points (although this has been accurately questioned by Chmarra et al. (Chmarra et al., 2008), pointing out the existence of a retraction movement in the correct *modus operandi*). Quality metrics such as end-product analysis and error count, although much more variable in their definition, are also considered basic for a correct determination of surgical level (Satava et al., 2003).

New metrics are proposed continuously as the means to acquire and process them become available. Their validation and study pose as key research aspects in the development of new objective assessment programs. Analysis of speed profile was studied by (Sokollik et al., 2004), with inconclusive results. Sinigaglia et al. identified acceleration of movements as a key factor for determining surgical expertise, studying their power spectra (Sinigaglia et al., 2005). A related parameter, motion smoothness, has been proposed and used by authors such as Stylopoulos et al. (Stylopoulos et al., 2004). Chmarra et al. proposed measuring the angular area and volume of the movements performed (Chmarra et al., 2010) and employed them for automatic detection of surgical skills. Overall, the clinical significance of these metrics has yet to be further determined, and thus thorough validation is required before being clinically adapted to training curricula.

3. Virtual Reality simulators for objective skills' assessment

Ethical concern on patient safety has led to a tendency to bring the training and assessment processes out of the OR as much as possible. Live animals and human cadavers are used as bench models, which generates a moral debate. Box trainers have also become popular training means, offering simple but key tasks to develop the necessary basic and advanced surgical skills. Examples on different box trainers can be found in (Rosser et al., 1997; Scott et al., 2000; Fichera et al., 2005). However, the real breakthrough came with the first Virtual Reality simulators, which allowed for controlled training and objective skills' assessment, on exercises ranging from simple tasks to complex laparoscopic procedures.

3.1 Advantages and limitations

Virtual Reality simulators offer some advantages that can add certain value to the training and assessment of surgical skills. They allow for training on controlled environments, and are always available for the trainee, without the need of a supervisor (thus reducing associated costs). They are ideal for monitoring a surgeon's learning curve, and offer a wide range of metrics which can be used for objective assessment, both efficiency and quality driven. More importantly, they deliver immediate constructive feedback of results and errors to the trainee, which some authors identify as basic in any effective training program (Issenberg et al., 2005).

However, some limitations have slowed down their clinical implantation (Lamata, 2006a). First, there are resource-derived constraints, such as trainees' loaded schedules, which leave them little time for practice; or the costs resulting from the expensive technologies behind the simulators. There are cases in which these advanced and sophisticated systems are available in the hospital but residents do not find the time or motivation to train their skills with them. Secondly, Virtual Reality environments show limitations in realism and interaction, which might not be critical for their didactic value, but are nevertheless of key importance to gain the acceptance of physicians. Thirdly, there are mentality-driven constraints, such as thinking of a surgical simulator as a videogame with no didactic value. Prior experience with videogames can also be a handicap when facing virtual simulators. It can even happen that such systems will lull oneself to a false sense of security, built on the development of incorrect habits while getting used to a virtual environment.

3.2 State of the art

Over the past fifteen years, virtual simulation has become a reference on the field of surgical training, with many attempts, some more successful than others, to develop, and most importantly, validate diverse models.

Surgical simulators can be classified according to the interventional procedures they are aimed for. Thus, we may find examples of arthroscopic simulators for knee and shoulders, as the InsightArthroVR (GMV Healthcare, Spain); cystoscopy and colonoscopy oriented, such as UroMentor and GIMentor respectively (Simbionix, Israel); intravascular simulators as CathSim (Immersion Medical, USA) and VIST-VR (Mentice, Sweden); and even focused on ophthalmological procedures, as the EYESi simulator (VRMagic, Germany).

In the field of laparoscopic surgery, we can find several well-positioned simulators on the market. As of this day, the principal laparoscopic simulators currently on the market can be found in Table 2. To further characterise each of them we can establish (1) whether tasks offered are basic, advanced (motor skill training) or complex (motor and cognitive training); (2) whether they offer realistic anatomic scenarios for procedures' simulation, and (3) whether they offer force feedback to the trainee.

LapMentor (Simbionix, Lod, Israel – Cleveland, USA) • Simple and advanced tasks, surgical procedures • Realistic scenarios • Force feedback	
LapSim (Surgical Science Ltd, Göteborg, Sweden) • Simple and advanced tasks, surgical procedures • Realistic scenarios • Optional force feedback	
MIST-VR (Mentice AB, Göteborg, Sweden) • Simple tasks • Non-realistic scenarios • Optional force feedback	
Promis (Haptica, Dublin, Ireland – Boston, USA) • Simple tasks, surgical procedures • Real scenarios (Hybrid Simulator) • Optional force feedback	
SIMENDO (DeltaTech, Delft, Netherlands) • Simple tasks • Non-realistic scenarios • No force feedback	

Table 2. Main simulators currently on the market

A final mention must be done to the many prototypes, which, without reaching commercial status, contribute to the validation and recognition of simulators as useful educational tools. Among the research prototypes we may mention some as Vesta (Tendick et al., 2000), Kalsruhe (Kühnapfel et al., 2000), GeRTiSS (Monserrat et al., 2004), and SINERGIA (Lamata et al., 2007), which will be thoroughly presented in Section 4.

3.3 Taxonomy of didactic resources for Virtual Reality simulation

Virtual Reality simulators can be conceived as training and evaluation means built using different didactic resources. These resources are classified into three main categories (Lamata et al., 2006b) based upon the extent to which simulators (1) emulate reality (*fidelity resources*); (2) exploit computer capabilities such as new ways of interaction and guidance (*teaching resources*); (3) measure performance and deliver feedback (*assessment resources*).

Regarding this taxonomy, there are three main directions in the design of a simulator, which can be taken independently or in a combined fashion (see Fig. 2): (1) the improvement of Virtual Reality technologies for providing a better fidelity, (2) the enhancement of simulation by augmenting the surgical scene for providing guidance, and (3) the development of evaluation metrics for giving a constructive feedback to the trainee. This framework will now be used to evaluate and compare existing simulators, and to address the development of an optimal solution by assessing the value of these didactic resources.

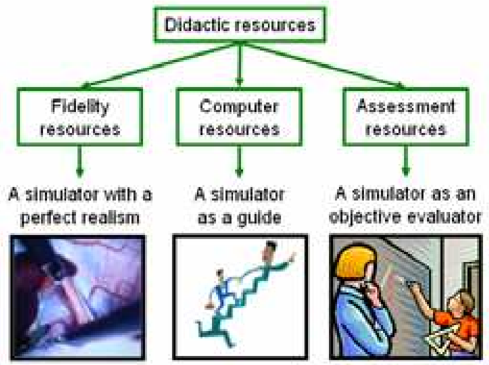

Fig. 2. The three conceptions of a Virtual Reality surgical simulator driven by the use of different didactic resources.

3.3.1 Evaluation and comparison of Virtual Reality simulators

Laparoscopic simulators, from simple box trainers with standardized tasks to advanced Virtual Reality simulators, are designed to train laparoscopic skills; but they make use of different didactic resources. A comparative analysis between some of the commercially available products is provided in Fig. 3.

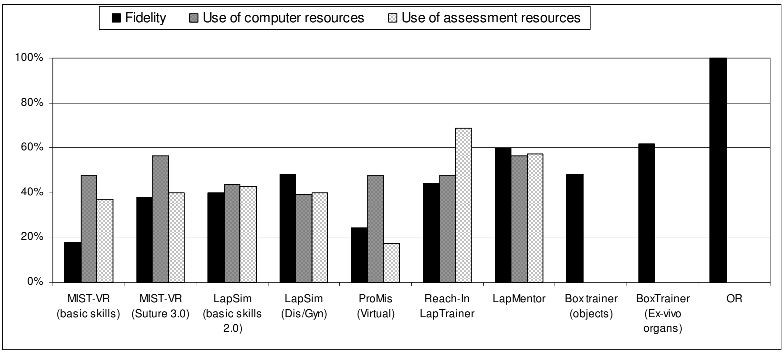

Fig. 3. Fidelity and use of computer and assessment resources by laparoscopic simulators.

Approaches to simulator design can be identified after studying how laparoscopic simulators make use of different didactic resources. The simplest one is an abstract conception of the surgical workspace focusing attention on the basic psychomotor skills that have to be developed by the trainee. MIST-VR "basic skills" was designed in this way, with an extremely simple interaction, almost no deformation and useful interaction indicators. The second approach aims at simulating a virtual patient with perfect realism, which is normally requested by surgeons. Force feedback is incorporated, organs are more realistic and interaction is enhanced. This is the trend usually followed by research institutions and companies, a trend lead by LapMentor as the simulator with the highest fidelity in almost every field (see Fig. 3).

But there is one last approach that might have a great potential: to enhance a simulator with a "virtual instructor" to guide the trainee through the procedure and deliver constructive feedback. Simulators make use of computer and assessment resources that build this "virtual instructor" capability. MIST-VR "Suture 3.0", which has the highest use of computer resources together with LapMentor (57%), offers an interesting guided interaction to teach trainees stitching and knotting skills. Reach-In Lap Trainer (nowadays integrated with MIST-VR), which had the highest use of assessment resources (69%), gave feedback about surgical performance not with low significant measurements like time or movements, whereas with what could be the advice of a surgical expert, with messages like "too much tissue bitten". The value of these types of resources has not yet been properly studied.

3.3.2 Towards an optimal design of a surgical simulator

Designing an optimum Virtual Reality surgical simulator for surgical training and assessment requires a suitable combination of Virtual Reality didactic resources. The value and importance of each of these didactic resources should therefore be assessed.

An important research question to be answered is to find the relationship between fidelity and training effectiveness. It would be really useful to assess how an increment in the realism of a simulation enhances or not the didactic capability. Fig. 4 shows a hypothetical line that relates these two variables for a given training objective, for example the acquisition of hand-eye coordination. The shape of this line is driven by three hypotheses. (1) "*A low degree of fidelity is enough to provide a good training effectiveness. It could even be the most efficient alternative*", based in the fact that skills acquired with a simple surgical simulator, MIST-VR, are transferred to the operating room (Seymour, 2002); (2) "*Incorporation of force feedback in simulation delivers an increase of training effectiveness in training*"; and (3) "*Stress present in real operating theatres decreases training effectiveness*". Several experiments are needed to figure out the real shape of this relationship between fidelity and training outcome.

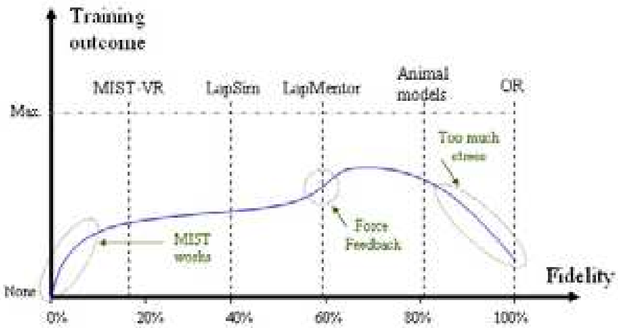

Fig. 4. Hypothetical relationship between simulation fidelity and training outcome. Fidelity values of commercial simulators are taken from (Lamata et al., 2006b).

On the other hand it is important to assess the value of computer and assessment resources offered. It could be contrasted if (1) "*Computer and assessment resources can overcome some lack of fidelity and result in an even more didactic simulator*", based on the fact that many times some interaction limitation is solved with a virtual interaction paradigm, for example when some colour code substitutes force feedback (Kitawaga et al., 2005). Other hypotheses are (2) "*Growing semitransparent spheres are a good forces substitute in suture training*"; (3) "*Suture training in Virtual Reality is enhanced with a guided training strategy focusing the fidelity resources* *on pre-defined ways of interaction compared to a non-guided one*"; (4) "*A guided training strategy with constructive feedback in Virtual Reality can enhance suture training outcome beyond that of physical trainers despite some fidelity limitations*"; or even (5) "*Computer and assessment resources can substitute an expert teacher behind the surgical trainee*".

3.4 Technical development of a Virtual Reality simulator

There are basically three main components in a Virtual Reality simulator: (1) a haptic interface to emulate the laparoscopic tools; (2) a monitor that simulates the abdominal cavity; and (3) a computer that manages both interfaces and runs the simulator's software, which in turn comprises four main modules (Fig. 5) (Lamata et al., 2007):

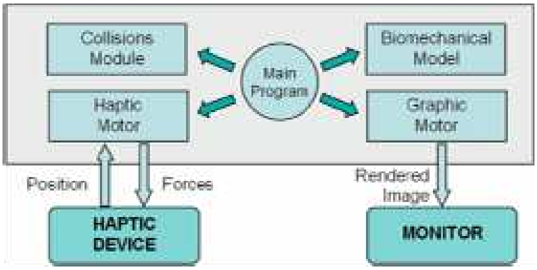

Fig. 5. Main Components in a Virtual Reality simulator

- **Biomechanical model:** Due to their very restrictive conditions, which imply update rates of at least 25Hz, robustness, satisfaction with visual result and precision, a trade-off must be achieved between complexity and speed when designing biomechanical models. There are two main approaches to be adopted (Meier et al, 2005): (1) heuristic models, (e.g. mass-spring models), or (2) models based on continuum mechanics, (e.g. finite elements models). There are several difficulties in this modelling process: biomechanical properties must be correctly acquired and tissue characteristics such as anisotropy, incompressibility and non-linearity considered (Picinbono et al., 2002). The models must represent surgical alterations that occur on real interventions, like cuts, dissections and other topological changes. Simplifications are made to address these problems, mainly assuming linear elasticity (valid for small deformations) and reducing the models' requirements. These have the drawback of being less realistic and prone to anomalous deformations (Picinbono et al., 2002).
- **Collisions' detection module:** Responsible of detecting overlapping objects and handling these detected collisions. There are three main types of collisions to manage: (1) tool-tool, (2) tool-tissue and (3) tissue-tissue. Implementation must consider the time constraints derived by fast-moving instruments. It is usually addressed with a coarse remodelling of the objects present in the scene to reduce their complexity. One simple alternative consists on defining boundary boxes on objects, and detecting any overlapping between them (Teschner et al, 2005). A recent advance introduces a fuzzy logic approach for handling tool-tissue collisions (García-Pérez et al., 2009).
- **Haptic rendering:** Delivering force feedback is still an unripe technology compared to visual rendering. This sensorial information requires a minimum update rate of 300Hz, which is technologically much more demanding compared to the 30Hz of visual display. There are several approaches for implementing it, like using the biomechanical

model (Delingette, 1998) or with a simplified geometric constraint force calculation proportional to the penetration depth of the tool (Balaniuk & Laugier, 2000).

- **Visual rendering:** Thanks to the many advances on computer graphics and the great deal of open source libraries available, visual rendering is a mature technology. Advances are focused now on photorealistic rendering and simulation of fumes and bleeding (Aggarwal et al., 2003).

3.5 Validation and acceptance of Virtual Reality simulators as assessment tools

As explained before, to be considered a suitable assessment tool, a measuring test or device and its related metrics must comply with a series of validation milestones. Virtual Reality simulators are not an exception to this, and so we can find in the literature many examples on the efforts to validate the different models available.

In the beginning, validation as a training means attracted much of the attention, focusing on concepts such as concurrent validity, skills transfer or the learning curve associated to the simulator (Lamata, 2006a). It is safe today to assume that Virtual Reality simulators are a valid supplementary method for surgical training, as effective as that provided by video-based box trainers. For further information, the reader is referred to (Gurusamy et al., 2009) for a complete meta-analysis of surgical simulators' validation for surgical training.

When it comes to validation as assessment tools however, there are more doubts about their reliability and fidelity (Thijssen & Schijven, 2010). For one, the limitations exposed previously still continue to hold sway among many clinicians. Indeed, validation studies up to today are sometimes inconclusive, and many surgeons are still mistrustful about their assessment capabilities.

3.5.1 Validation strategies

Different strategies are employed to carry out the validation studies necessary for a simulator (Fried & Feldman, 2008). Face and content validation, being for the most part subjective studies, are usually done by means of structured questionnaires and reviews. Face validity questionnaires usually call for personal opinions on the simulator's usefulness "at face value"; whilst content validation requires a more thorough and complete review of the tasks, skills assessed and metrics employed by the expert reviewer, before passing judgment.

Construct validation is usually granted if the simulator is able to determine differences between groups of surgeons with known different skill levels (as for example, residents and expert surgeons). The strategy employed is to divide the test population according to these levels, and measure their performance and the differences observed on the simulator. Some factors may however dampen the results of a study if not properly considered. Test subjects should not have prior experience with the simulator, as their learning curve may be enhanced because of this. Also, some studies have shown the influence of video-gaming experience as an influential concern (McDougall et al., 2006). Well designed studies, as well as increasing the test population, may help to mitigate these issues.

Concurrent and predictive validations require an alternative and valid assessment method to be deployed (a *gold standard*), in order to compare the results obtained on both scenarios. Test subjects are usually grouped by levels, and similar tasks performed and measured in both settings. Finding a gold standard for this comparison is not always easy; usually, comparison is done via OSATS, motion tracking systems, or employing another virtual simulator. If the comparison is done in a similar time period, it is considered concurrent validation; if on the contrary the time lapse between them is considerable, it is predictive validation which is being measured.

3.5.2 Validation studies on commercial simulators

There are many reports considering assessment validation of surgical simulators in the literature, especially where construct and concurrent validity are concerned. This is reasonable if we consider that these two parameters are essential for the automatic and immediate assessment they intend to carry out. Table 3 briefly summarizes the conclusions arrived at on the main studies performed on commercial simulators for the last few years.

Simulator as a skill assessment tool		
Face validity		
Simulator	***Reference***	***Valid?***
MIST-VR	(Maithel et al., 2006)	Yes
LapSim	(Schreuder et al, 2009)	Yes
SIMENDO	(Verdasdoonk et al., 2006)	Yes
LapMentor	(McDougall et al., 2006)	Yes
	(Ayodeji et al., 2007)	Yes
ProMis	(Botden et al., 2008)	Yes
Content validity		
Simulator	***Reference***	***Valid?***
SIMENDO	(Verdasdoonk et al., 2006)	Yes
LapMentor	(McDougall et al., 2006)	Yes
Construct validity		
Simulator	***Reference***	***Valid?***
MIST-VR	(Gallagher et al., 2004)	Yes
LapSim	(Sherman et al., 2005)	Yes
	(Ro et al., 2005)	No at the first exposure to simulator
	(Langelotz et al., 2005)	Yes, but time and path metrics only
	(Hassan et al., 2005)	Yes, more patent in second
	(Eriksen & Grantcharov, 2005)	Yes
	(Woodrum et al., 2006)	Yes, but only some parameters
	(Larsen et al., 2006)	Yes
	(Schreuder et al., 2009)	Yes
SIMENDO	(Verdasdoonk et al., 2007)	Yes
LapMentor	(McDougall et al., 2006)	Yes
	(Zhang et al., 2008)	Yes
	(Aggarwal et al., 2009)	Yes
ProMIS	(Broe et al., 2006)	Yes
	(Neary et al., 2007)	Yes
	(Pellen et al., 2009)	Yes
Concurrent validity		
Simulator	***Reference***	***Valid? / Concurrent with…?***
MIST-VR	(Gallagher et al., 2004)	A little, with OR metrics
LapSim	(Youngblood et al., 2005)	Yes, with box trainer
	(Newmark et al., 2007)	Yes, with box trainer
SIMENDO	(Verdasdoonk et al., 2006)	Yes
LapMentor	(Okrainec et al., 2008)	Yes, with GOALS metrics in the OR
ProMis	(Ritter et al., 2007)	Yes, with OR metrics
	(Botden et al., 2007)	Yes, with LapSim
Predictive validity		
Simulator	***Reference***	***Valid?***
LapSim	(Hassan et al., 2008)	Yes
LapMentor	(Greco et al., 2008)	Yes
ProMis	(McCluney et al., 2006)	Yes

Table 3. Validation studies on commercial Virtual Reality simulators for skills' assessment

As promising as these results are, they are only but the first milestone on the slow road to integrating simulators on the design of structured assessment curricula.

4. Designing a Virtual Reality simulator for surgical training and assessment: The SINERGIA experience

There is little specific literature about how to develop an efficient didactic design for a simulator. It can be found that an ergonomic task analysis (Stone & McCloy, 2004) was used for the design of the MIST-VR, but without any further detail. The construction of the SINERGIA laparoscopic Virtual Reality simulator is one of the best documented examples of the development process (Lamata, 2006a), and this section highlights its main aspects. For a more detailed and thorough description of the design and development process of surgical systems the reader is referred to (Freudenthal et al., 2010).

4.1 Didactic design of simulator tasks

Design of the didactic contents of a simulator is based on a thorough analysis of the training needs, driven by a surgical training curriculum. Existing solutions and validation studies are also an important reference for the definition of specifications, which are described with a suitable use of simulation technologies. The third main pillar in the designing process is understanding the capabilities and reach of Virtual Reality technologies.

Human beings have perceptual limitations of the sensory, motor and cognitive system. Laparoscopy is characterised by a loss of sensory stimuli of the surgeon, which leads to the need of developing new skills. Knowing and understanding how surgeons interact in the surgical theatre and develop their skills is an important issue in order to address the design of a surgical simulator. This contributes to the definition of the required degree of simulation fidelity, a very controversial issue. For example, it is unclear the role of force feedback in surgical training (Kneebone, 2003). Comprehension of the laparoscopic interaction leads also to the definition of objective metrics of surgical skill. For example, an analysis of tissue consistency perception (Lamata et al, 2006c) led to the definition of “Force sensitivity” training tasks in the SINERGIA simulator (Lamata et al., 2007).

Training objectives and needs of the SINERGIA laparoscopic simulator were grounded on the vast training experience of the Minimally Invasive Surgery Centre Jesús Usón (MISCJU, Cáceres, Spain). This Centre has a thoroughly validated methodology of training based on four levels: (1) basic and advanced skills with box trainers, (2) anatomical protocols and advanced skills with animal models, (3) advanced procedural skills with tele-surgical applications and (4) practice in the OR (Fig. 6).

The SINERGIA laparoscopic simulator was conceived as a means for training and assessment on the first level in the pyramidal model. An analysis of the laparoscopic skills acquired at this stage led to the definition of seven didactic units: hand–eye coordination, camera navigation, grasping, pulling, cutting, dissection and suture.

4.2 Technical development

SINERGIA was developed in C++ language, with WTK libraries (WorldToolKit, Engineering Animation Inc., Mill Valley, CA-USA) and in a Windows environment. The chosen haptic interface was the Laparoscopic Surgical Workstation (Immersion Medical, Gaithersburg, USA) (Lamata et al., 2007).

Fig. 6. MISCJU Formation Pyramid

Biomechanical modelling was solved employing T2-Mesh mass spring models for hollow objects. A T2-Mesh model is a surface model that seeks for simplicity and speediness of calculi, in detriment of a realistic behaviour. Nodes in the model have a mass assigned to them. They are linked with linear springs which act as energy storage and react against deformations. The equations' system is relatively small and, therefore, fast in its resolution. Nevertheless it is an iterative model and consequently mined by the risk of instabilities and oscillations (Meier et al., 2005). Solid object handling was improved by the use of ParSys, which are fast and stable models composed by a set of interconnected volumetric elements, also called particles. Volumetric behaviour is given and guaranteed by its structure and volume conservation by the constant number of particles in an object (Pithioux et al., 2005).

Collision detection was performed by an *ad hoc* library built for SINERGIA, which tests geometrically the interaction between tools (rigid objects) and deformable objects, and manages topological changes. Assuming that the objects are modelled by means of a triangular mesh, **collision handling** is posed as finding the new positions of the vertices of the triangles involved in the collision. The problem is based on the tool kinematics (the tool velocity vector) and the normal vectors to the triangles involved in the collision. Therefore, each of its vertexes are displaced out of the tool bearing in mind the fuzzy nature of the tool motion, which is modelled as penetration or sliding (García-Pérez, 2009).

Graphical design of the scenarios was done with Blender (Blender Foundation, Amsterdam, Netherlands), an open source and multi-platform tool for 3D design and modelling, with a python-command interface. It provided the advantages of the ability to share 3D scenes from different platforms, and the feasibility to share models generated from real medical images. It was employed to design scenes (including exercises) and add the effects of texture and realism. Fig. 7 shows an example of a pig's abdominal cavity built using these models.

Fig. 7. Example of a surgical scene

4.3 SINERGIA objective assessment module

Addressing the needs described for surgical assessment, the SINERGIA Virtual Reality simulator includes an objective evaluation system, in order to monitor trainees' learning curves (Fig. 8). Objective metrics' definition allows trainees to learn from their mistakes by means of indications when errors are performed (*formative feedback*) or by visualization of the global practice score (*summative feedback*).

This evaluation component can be used by three different groups: (1) trainees, who perform the tasks and whose metrics are stored and managed by the system; (2) teachers, who monitor and follow-up the trainees' progress by means of the evaluation interface; and (3) administrators, tasked with system and user management. Security is an important issue in the evaluation system. Thus, a teacher is only able to follow his pupils' results, and a trainee tracks only his own data, but can compare these results with an average mark of the global users' community.

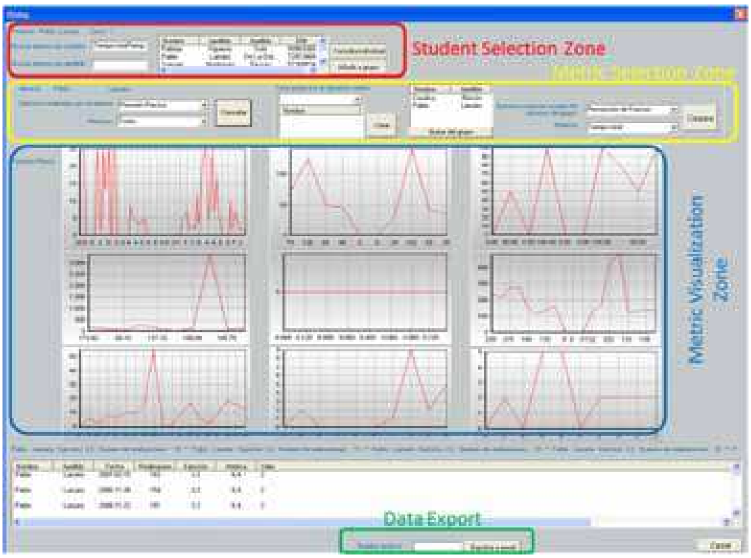

Fig. 8. SINERGIA Assessment Module

In order to manage all evaluation data, different graphics modalities are implemented for easy monitoring and understanding of the surgical skills' evolution of the trainee. Comparisons between different individuals or groups of pupils and data exchange for statistical analysis are possible in the user interface.
Moreover, the system offers an easy to use interface which allows efficient metric management, while dealing with huge amounts of information. Its design has been validated by expert surgeons of the Minimally Invasive Surgery Centre Jesús Usón.

4.4 SINERGIA face and construct validation

First validation of SINERGIA consisted on two tests to determine face and construct validity (Sánchez-Peralta et al., 2010). Among all tasks provided, five were selected for the study (hand-eye coordination, camera navigation, navigation and touch, accurate grasping and coordinated pulling, as shown in Fig. 9).

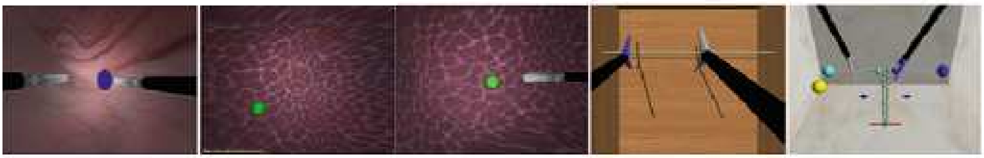

Fig. 9. Tasks performed in the validation process: From left to right: hand-eye coordination, camera navigation, navigation and touch, accurate grasping and coordinated pulling.

10 novices and 6 expert surgeons took part in the validation experiment. Each subject performed each task in the SINERGIA Virtual Reality simulator once and filled in demographic and face validity questionnaires after performance. No external help was given during the exercises; only a brief explanation of the task prior performance. Results are shown in Table 4.
Significance was calculated using the Mann-Whitney U test for $P<0.05$. Statistical analysis highlighted significant differences between the experienced and non-experienced groups in 60% of the evaluated metrics, which implies a partial construct validity being reached. Face validity is confirmed in the questionnaires, where there are no significant differences in any of the evaluated aspects. Results also showed that both groups considered the most remarkable characteristic in SINERGIA its usefulness as a learning tool for basic laparoscopic skills, rating it with the highest possible score.

5. Alternative technologies for objective surgical evaluation

In order to give a complete picture of the current state of the art on surgical assessment technologies, and to understand where Virtual Reality simulators stand in the bigger picture, we will briefly present other ways for acquiring efficiency measurements for objective evaluation.
On the last few years several systems have been developed for force and movement analysis (Moorthy et al, 2003). All these have in common the need for some means to capture the value of objective metrics. This is usually achieved by active or passive tracking devices. Active tracking relies on optical, electromagnetic or mechanical sensors mounted on the surgical tools. Passive tracking relies on the external detection of markers placed on the instruments, using ultrasound or electromagnetic technologies. The reader is referred to (Chmarra et al., 2007) for more information behind the technologies behind tracking systems.

SINERGIA construct validity				
Tasks	**Metrics**	**Novices (n = 14)**	**Experts (n = 6)**	**P value**
Coordination	Total time (s)	75.16 ± 9.72	61.97 ± 11.11	0.033
	Partial time (s)	2.98 ± 0.49	2.48 ± 0.45	0.062
	Fulfilment (%)	75.14 ± 8.18	85.33 ± 8.26	0.051
	L-I efficiency (%)	36.80 ± 11.48	46.97 ± 10.67	0.062
	R-I efficiency (%)	37.06 ± 8.48	47.27 ± 10.78	0.033
	Harms to background (#)	11.43 ± 5.45	4.67 ± 2.88	0.006
Navigation	Total time (s)	104.71 ± 10.95	97.50 ± 12.93	0.353
	Partial time (s)	7.86 ± 0.73	7.36 ± 0.57	0.207
	Fulfilment (%)	76.36 ± 14.48	88.00 ± 8.39	0.076
	L-I efficiency (%)	40.93 ± 8.19	47.52 ± 3.89	0.033
	Harms to background (#)	0.29 ± 0.61	0.67 ± 0.82	0.353
Navigation and touch	Total time (s)	106.79 ± 19.92	85.33 ± 11.36	0.005
	Partial time (s)	7.48 ± 1.27	5.97 ± 0.92	0.007
	Fulfilment (%)	71.64 ± 15.83	95.50 ± 4.93	0.007
	L-I efficiency (%)	40.93 ± 10.03	55.45 ± 3.84	0.014
	Harms to background (#)	95.36 ± 94.01	11.33 ± 4.46	0.002
Precise grasping	Total time (s)	50.14 ± 12.66	32.50 ± 5.58	0.002
	Partial time (s)	5.01 ± 1.30	3.27 ± 0.52	0.002
	Fulfilment (%)	100.00 ± 0.00	100 ± 0.00	1
	Deviation from central point (cm)	0.06 ± 0.01	0.04 ± 0.01	0.003
	L-I efficiency (%)	6.43 ± 2.62	8.33 ± 2.16	0.091
	R-I efficiency (%)	8.21 ± 2.75	10.17 ± 3.54	0.207
	Grasps out of the area (#)	4.71 ± 6.39	1.00 ± 1.26	0.02
	Grasps with excessive pressure (#)	3.00 ± 2.08	0.00 ± 0.00	0.002
Coordinate Traction	Total time (s)	123.71 ± 45.41	87.00 ± 24.58	0.051
	Partial time (s)	41.33 ± 15.24	29.11 ± 8.32	0.062
	Fulfilment (%)	69.05 ± 20.52	94.44 ± 13.61	0.026
	L-I distance from ideal line (cm)	836.93 ± 352.73	501.33 ± 201.78	0.02
	R-I distance from ideal line (cm)	748 ± 285.64	504.67 ± 184.19	0.041
	L-I efficiency (%)	4.31 ± 1.03	6.61 ± 1.42	0.002
	R-I efficiency (%)	5.10 ± 1.66	7.11 ± 2.22	0.062
	Non-coordination moments (#)	31.64 ± 36.56	1.33 ± 3.27	0.001

Table 4. Metrics results for novices and experts

The technology behind these devices is all but the same than that employed by the haptic systems of Virtual Reality simulators (indeed, they can be exchangeable), so the main difference resides in the tracking system's application: instead of software-based virtual tasks, they are used as training and assessment means on box trainers and even the OR.

Force sensing has been mainly approached by Rosen et al. (Rosen et al., 2002). They demonstrated that experienced surgeons apply higher force/torque magnitudes during tissue dissection than novices, and vice versa for tissue manipulation. Sensing was performed by a specially built system, the BlueDragon, a bulky device with built-in mechanical sensors. More recently, Horeman et al. have approached tool/tissue forces detection by means of a pressure platform placed under the box trainer task (Horeman et al., 2010).

However, movement sensing has been one of the most common approaches followed. Motion tracking systems register both the position (x,y,z coordinates) and orientation (yaw, pitch and roll) of the surgical tools. Systems have been developed that place position sensors

on tools or on hands and fingertips, such as ADEPT (Hanna et al., 1998), ICSAD (Datta et al., 2001), CELTS (Stylopoulos et al., 2004), or TrEndo (Chmarra et al., 2006). Among other things, measuring trajectories has been used for speed and acceleration calculi, optimal path and economy of movement determination, depth perception and movement sequences analysis, and repetitions or idle states detection. Active sensing has the disadvantages of introducing new elements on the surgical theatre, thus altering it, and also of modifying the tools' ergonomics. Passive sensing systems have thus been also developed for motion tracking, such as the Zebris (Sokollik et al., 2004).
An interesting approach for passive sensing can be the analysis of the laparoscopic videos, which allow tracking of movements employing computer vision techniques. This way, information about position and trajectories can be acquired, which can then be used to obtain speeds and accelerations, and for metrics calculation such as economy of movements, efficiency, optimal path, etc. The challenge of this approach is to exploit the 2D information of the surgical scenario captured by the endoscope in order to assess the laparoscopic tool's 3D position. This approach solves the problem of calculating the 3D position and orientation of a tool with only the 2D information extracted from each frame of a video sequence. Combining segmentation and edge detection techniques, position of the surgical tools' borders is determined. Knowing the tool's cylindrical geometrical dimensions and its 2D projection as denoted by the detected borders, real 3D tool's pose is calculated (Fig. 10). The mathematical equation for this calculus is a description of the geometrical relations between the tools, trocars and the optical centre of the camera; its complete description and explanation can be found in (Cano et al., 2008). Current tracking performance of these methods, with an accuracy of 9.28mm, is good enough for gesture analysis and objective evaluation of surgical manoeuvres.

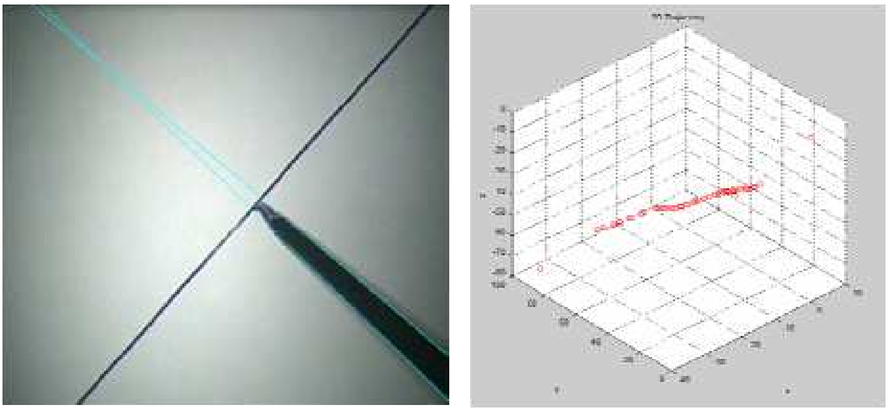

Fig. 10. Video tool tracking: Left, 2D detection of the instrument's borders; right, 3D trajectory determination

6. Discussion

Developing structured training curricula for laparoscopic surgery has become a priority in the past few decades. There are several reasons behind this: primarily, there is an ever-growing social concern on patient safety, whether on medical errors or on the ethics behind training on real patients. It has become also necessary to optimize resident's timetables, and evolve efficient training programs around them. Finally, there is a need to maximize the efficiency/cost ratios of these programs.

Within these new curricula, Virtual Reality simulators, along with the aforementioned tracking systems, present themselves as useful methods for training and assessing skills in laboratory settings. In these conditions, the trainee is able to practice his skills in a stress-free environment and receive objective feedback of his technical performance. A good example on trying to integrate Virtual Reality simulators on surgical curricula can be found in (Aggarwal et al., 2009).

However, controversy has surrounded their use since the very first models. Their lack of realism has been one of the strongest arguments employed against them. Much work is still required in this field; some authors have even pointed out that surgeons and trainees seem to prefer training on physical simulators due to the more realistic visual and tactile sensations involved (Chmarra et al., 2008; Gurusamy et al., 2009). Validation studies have, however, proved over the last few years that these limitations do not diminish the training effectiveness; indeed, where the ultimate goal is the acquisition of motor skills such as hand-eye coordination or depth perception, a realistic scenario does not necessarily add much value to the training process.

But whilst training effectiveness seems to be generally accepted, the same cannot still be said of their assessment capabilities (Thijssen & Schijven, 2010). This, however, can be partly blamed on the difficulties of defining a standardized training curriculum, and its associated metrics. Although this problem could be extended to other areas of surgical training besides Virtual Reality simulators, in their case it becomes magnified if considered along their limitations and their slow acceptance (Liselotte & Dewan, 2009).

The truth is that, no matter what, the possibilities of Virtual Reality simulators as evaluation tools are potentially great. The fact that they are ever available for training, their reproducibility and the immediate feedback they provide mean that training and assessment of skills can be easily accommodated to the trainee's schedule. Also, their unrivalled ability to capture and acquire not only efficiency metrics but also quality based ones gives them an important advantage over other acquisition devices. The variety of tasks available, from simple exercises to complex interventions, allows also extending their range to the field of cognitive knowledge.

No doubt in the future Virtual Reality simulation will focus on the improvement of the visual and tactic experience for the user; as technologies become available. But it will be interesting to see their clinical evolution concerning their assessment capabilities. One promising research area is the automatic determination of surgical level. These techniques have already been explored by some authors (Rosen et al., 2002, Chmarra et al., 2009; Megali et al., 2006), applying techniques such as Hidden Markov Models or Linear Discriminant Analysis. Their inclusion on future generations of Virtual Reality simulators could help improve their value as objective assessment tools.

To conclude, it is necessary to stress the complimentary role that Virtual Reality simulators and other tracking devices play in the much larger scope of the surgical structured curricula. Despite the need for objectivity, the fact remains that final expertise accreditation should come by the hand of an expert mentor. There will always be a subjective component to the determination of a surgeon's readiness, which will imply judging abilities such as reaction time, mentality, patient care, handling of stress or group working capability. These are all important factors that add a human dimension to the qualification process, and thus should always be considered.

7. Conclusion

Studies on Virtual Reality simulators over the last few years have focused especially on their training capabilities, determining whether motor skills acquisition on them is really effective and if they really translate to the OR afterwards. The present chapter has presented them from a different, although related, perspective: their usefulness as skills' assessment tools. While acceptance of Virtual Reality simulators is growing by the day as validation studies prove their usefulness, the development of training curricula to determine surgical expertise is not being so easy. It has been the authors' goals in this chapter to convey to the reader the actual state of the art of Virtual Reality simulators in the field of skills assessment, in relation to other devices such as sensor-based tracking systems; to point out to him their limitations and advantages; and further still, to give him insight on the development and validation process of a surgical simulator (SINERGIA) in order to prove that, despite their limitations and the complications surrounding them, we believe that in the near future, Virtual Reality will play an important role on structured and objective skills assessment.

8. References

Aggarwal, R.; Bhasin, Y.; Raghupathi, L. & Devarajan, V. (2003). Special visual effects for surgical simulation: Cauterization, irrigation and suction. *Medicine Meets Virtual Reality 11*, Nextmed: Health Horizon, 94, pp. 1-3.

Aggarwal, R. ; Crochet, P.; Dias, A.; Misra, A.; Ziprin, P. & Darzi, A. (2009). Development of a virtual reality training curriculum for laparoscopic cholecystectomy. *British Journal of Surgery*, 96 pp. 1086–1093

Ayodeji, I.D.; Schijven, M.; Jakimowicz, J. & Greve, J.W. (2007). Face validation of the Simbionix LAP Mentor virtual reality training module and its applicability in the surgical curriculum. *Surgical endoscopy*, 21, 9, pp. 1641-1649.

Balaniuk, R. Y. & Laugier, C. (2000). Haptic interfaces in generic virtual reality systems, Proc. *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2000, (IROS 2000)*, Takamatsu Japan, 2, pp. 1310-1315.

Bloom, B.; Englehart, M.; Furst, M.; Hill, W. & Krathwohl, D. Taxonomy of educational objectives: The classification of educational goals. New York: David McKay Company, Inc.; 1956.

Botden, S.M.B.I.; Buzink, S.N.; Schijven, M.P. & Jakimowicz, J.J. (2007). Augmented versus Virtual Reality Laparoscopic Simulation: What is the difference? *World journal of surgery*, 31, 4, pp. 764-772.

Botden, S.M.; Berlage, J.T.; Schijven, M.P. & Jakimowicz, J.J. (2008). Face Validity Study of the ProMIS Augmented Reality Laparoscopic Suturing Simulator. *Surgical technology international*, 17, pp. 26-32.

Broe, D.; Ridgway, P.F.; Johnson, S.; Tierney, S. & Conlon, K.C. (2006). Construct validation of a novel hybrid surgical simulator. *Surgical endoscopy*, 20, 6, pp. 900-904.

Cano, A.M.; Gayá, F.; Lamata, P.; Sánchez-González, P & Gómez, E.J. (2008). Laparoscopic tool tracking method for augmented reality surgical applications. Lecture Notes in Computer Science, vol. 5104, pp. 191-196, 978-3-540-70521-5_21.

Cavallo, F.; Megali, G.; Sinigaglia, S.; Tonet, O. & Dario, P. (2005). A Biomechanical Analysis of Surgeon's Gesture in a Laparoscopic Virtual Scenario. *Studies in health technology and informatics*, 119, pp. 79.

Chmarra, M.K.; Bakker, N.H.; Grimbergen, C.A. & Dankelman, J. (2006). TrEndo, a device for tracking minimally invasive surgical instruments in training set-ups. *Sensors and Actuators*, 126, pp. 328-334

Chmarra, M.K.; Grimbergen, C.A. & Dankelman, J. (2007). Systems for tracking minimally invasive surgical instruments. *Minimally Invasive Therapy & Allied Technologies*, 16, 6, pp. 328-340

Chmarra, M.K.; Dankelman, J.; Van Del Dobbensten, J.J. & Jansen, F.W. (2008). Force feedback and basic laparoscopic skills, *Surgical Endoscopy*, 22, pp. 2140–2148.

Chmarra, M.K.; Klein, S.; de Winter, J.C.F.; Jansen, F.W. & Dankelman, J. (2010). Objective classification of residents based on their psychomotor laparoscopic skills. *Surgical Endoscopy*, 24, 5, pp. 1031-1039.

Cotin, S.; Stylopoulos, N.; Ottensmeyer, M.; Neumann, P.; Rattner, D. & Dawson, S. (2002). Metrics for Laparoscopic Skills Trainers: The Weakest Link! *Proceedings of Medical Image Computing and Computer-Assisted Intervention—MICCAI 2002*, Tokyo, Japan, September 2002.

Cuschieri, A., (2005). Laparoscopic surgery: current status, issues and future developments. *Surgeon*, 3, 3, 125-138.

Datta; V., Mackay, S., Mandalia, M. & Darzi, A. (2001). The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *Journal of the American College of Surgeons*, 193, 5, pp. 479-485.

Datta, V.; Mandalia, M.; Mackay, S.; Chang, A.; Cheshire, N. & Darzi, A. (2002). Relationship between skill and outcome in the laboratory-based model, *Surgery*, 131, 3, pp. 318-323.

Delingette, H. (1998). Toward realistic soft-tissue modeling in medical simulation, *Proceedings of the IEEE*, 86, 3, pp. 512-523.

Eriksen, J.R. & Grantcharov T. (2005) Objective assessment of laparoscopic skills using a virtual reality stimulator, *Surg. Endosc.*, 19, 9, pp. 1216-1219.

Fichera, A.; Prachand, V.; Kives, S.; Levine, R. & Hasson, H. (2005). Physical reality simulation for training of laparoscopists in the 21st century. A multispecialty, multi-institutional study. *Journal of the Society of Laparoendscopic Surgeons*, 9, 2, pp. 125-129.

Freudenthal, A.; Stüdeli, T.; Lamata, P. & Samset ,E (2010). Collaborative co-design of emerging multi-technologies for surgery. *Journal of Biomedical Informatics* (in press)

Fried, G.M. & Feldman, L.S. (2008). Objective Assessment of Technical Performance. *World journal of surgery*, 32, 2, pp. 156-160.

Gallagher A.G.; Ritter E.M. & Satava R.M.. (2003). Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surgical Endoscopy*, 17, pp. 1525-1529.

Gallagher, A.; Lederman, A.; Mcglade, K.; Satava, R. & Smith, C. (2004). Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance. *Surgical endoscopy*, 18, 4, pp. 660-665.

García-Pérez, V.; Muñoz-Moreno, E.; Aja-Fernández, S. & Alberola-López, C. (2009). A 3-D Collision Handling Algorithm for Surgery Simulation Based on Feedback Fuzzy Logic. *IEEE Transactions on Information Technology in Biomedicine*, 13, 4, pp. 451-457.

Greco, E. M.D.; Escallon, J. M.D.; Grantcharov, T. M.D.; Okrainec, A. M.D. (2008). Do Performance Measures on the LapMentor VR Simulator Predict FLS Performance? *In Canadian Association of General Surgeons Annual Meeting*, September 2008, Halifax, Nova Scotia, Canada.

Gurusamy, K. S.; Aggarwal, R.; Palanivelu, L. & Davidson, B. R. (2009). Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database of Systematic Reviews.*, Jan 21, 1, CD006575.

Halsted, W.S. (1904). The training of the surgeon. *Bulletin of the Johns Hopkins Hospital* 15, pp. 267-275.

Hanna, G.; Drew, T.;Clinch, P.; Hunter, B. & Cuschieri, A. (1998). Computer-controlled endoscopic performance assessment system. *Surgical endoscopy*, 12, 7, pp. 997-1000.

Harden, R.M.; Stevenson, M.; Downie, W.W. & Wilson, G.M. (1975). Assessment of clinical competence using objective structured examination. *British Medical Journal*, British Medical Association, 1, 5955, pp. 447-451.

Hassan, I.; Sitter, H.; Schlosser, K.; Zielke, A.; Rothmund, M. & Gerdes B. (2005). A virtual reality simulator for objective assessment of surgeons' laparoscopic skill. *Chirurg*, 76, 2, pp. 151-156.

Hassan, I.; Osei-Agymang, T.; Radu, D.; Gerdes, B.; Rothmund, M. & Fernandez, E.D. (2008). Simulation of laparoscopic surgery--four years' experience at the Department of Surgery of the University Hospital Marburg. *Wiener klinische Wochenschrift*, 120, 3-4, pp. 70-76.

Horeman, T.; Rodrigues, S.P.; Jansen, F.W.; Dankelman, J. & van den Dobbelsteen, J.J. (2010). Force measurement platform for training and assessment of laparoscopic skills. *Surgical Endoscopy*. Published online.

Issenberg, S.B.; McGaghie, W.C.; Petrusa, E.R.; Lee, G.D. & Scalese, R.J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*, 27, 1, pp. 10-28.

Kitagawa, M.; Dokko, D.; Okamura, A.M. & Yuh D.D. (2005). Effect of sensory substitution on suture-manipulation forces for robotic surgical systems. *The Journal of Thoracic and Cardiovascular Surgery*, 129, 1, pp. 151-158.

Kneebone R. (2003). Simulation in surgical training: educational issues and practical implications. *Medical Education*, 37, 3, pp. 267-277.

Kühnapfel, U.; C. Akmak, H. K. & Maass, H. (2000). Endoscopic surgery training using virtual reality and deformable tissue simulation. *Computers & Graphics*, 24, pp. 671-682.

Lamata, P., (2006). *Methodologies for the Analysis, design and evaluation of laparoscopic surgical simulators*. Similar Network of Excellence, UCL Presses, Universitaires de Louvain. Available at www.i6doc.com.

Lamata, P.; Gómez, E. J.; Bello, F.; Kneebone, R. L.; Aggarwal, R. & Lamata, F. (2006). Conceptual framework for laparoscopic VR simulators. *IEEE Computer Graphics and Applications.*, Nov-Dec, 26, 6, pp. 69-79.

Lamata, P.; Gómez. E.J.; Sánchez-Margallo, F.M.; Lamata, F.; del Pozo; F. & Usón, J. (2006). Tissue consistency perception in laparoscopy to define the level of fidelity in virtual reality simulation. *Surgical Endoscopy*, 20, 9, pp. 1368-75.

Lamata, P.; Gómez, E. J.; Sánchez-Margallo, F. M.; López, O.; Monserrat, C.; García, V.; Alberola, C.; Florido, M. A.; Ruiz, J. & Usón, J. (2007). SINERGIA laparoscopic virtual reality simulator: didactic design and technical development. *Computer Methods and Programs in Biomedicine*, 85, 3, pp. 273-83.

Langelotz, C.; Kilian, M.; Paul, C.; & Schwenk W. (2005). LapSim virtual reality laparoscopic simulator reflects clinical experience in German surgeons. *Langenbecks Arch. Surg*, 390, 6, pp. 534-537.

Larsen, C.R.; Grantcharov, T.; Aggarwal, R.; Tully, A.; Sørensen, J.L.; Dalsgaard, T. & Ottesen, B. (2006). Objective assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. *Surgical endoscopy*, 20, 9, pp. 1460-1466.

Liselotte, L.M. & Dewan, P. (2009). Virtual Reality simulators in gynecological surgery: A surging new wave. *Journal of the Society of Laparoendscopic Surgeons*, 13 pp. 279–286.

Maithel, S.; Sierra, R.; Korndorffer, J.; Neumann, P.; Dawson, S.; Callery, M.; Jones, D. and Scott D. (2006). Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? Surgical Endoscopy, 20,1, pp. 104-112.

Martin, J.; Regehr, G.; Reznick, R.; Macrae, H.; Murnaghan, J.; Hutchison, C. & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery*, 84,2, pp.273-278.

McCluney, A.; Feldman, L. & Fried, G. (2006). Validation of the ProMIS Hybrid Simulator using a standard set of laparoscopic tasks. *Proc.Sages*.

McDougall, E.M.; Corica, F.A.; Boker, J.R.; Sala, L.G.; Stoliar, G.; Borin, J.F.; Chu, F.T. & Clayman, R.V. (2006). Construct Validity Testing of a Laparoscopic Surgical Simulator. *Journal of the American College of Surgeons*, 202, 5, pp. 779-787.

Meier, U.; Lopez, O.; Monserrat, C.; Juan, M. C. & Alcaniz, M. (2005). Real-time deformable models for surgery simulation: a survey. *Computer Methods and Programs in Biomedicine*, 77, 3, pp. 183-197.

Miller, G.E. (1990). The assessment of clinical skills/competence/performance. Academic Medicine, 65, 9, pp. S63-S67.

Monserrat, C.; López, O.; Meier, U.; Alcañiz, M.; Juan, C. & Grau, V. (2003). GeRTiSS: A Generic Multi-model Surgery Simulation, *Proc. IS4TM*, Juan-Les-Pinnes (Fr), LNCS, 2673, 59-66.

Moorthy, K.; Munz, Y.; Sarker, S.K. & Darzi, A. (2003). Objective assessment of technical skills in surgery. *British Medical Journal*, 327, 7422, pp. 1032.

Neary, P.C.; Boyle, E.; Delaney, C.P.; Senagore, A.J.; Keane, F.B.V. & Gallagher, A.G. (2008). Construct validation of a novel hybrid virtual-reality simulator for training and assessing laparoscopic colectomy; results from the first course for experienced senior laparoscopic surgeons. *Surgical endoscopy*, 22, 10, pp. 2301-2309.

Newmark, J.; Dandolu, V.; Milner, R.; Grewal, H.; Harbison, S. & Hernandez, E. (2007). Correlating virtual reality and box trainer tasks in the assessment of laparoscopic surgical skills. *American Journal of Obstetrics and Gynecology*, 197, 5, pp. 546-546.

Okrainec, A.; Greco, E.; Arts, M.A. & Grantcharov, T. (2008), Does assessment during virtual surgery predict actual intraoperative performance? *Association for Surgical Education Annual Meeting*, Toronto, Canada.

Pellen, M.G.C.; Horgan, L.F.; Barton, J.R. & Attwood, S.E. (2009). Construct validity of the ProMIS laparoscopic simulator. *Surgical endoscopy*, 23, 1, pp. 130-139.

Picinbono, G.; Delingette, H. & Ayache, N. (2002). Non-linear elastic deformable models for real-time surgery simulation. *Comptes rendus biologies*, 325, 4, pp. 335-344.

Pithioux, M.; Lopez, O.; Meier, U.; Monserrat, C.; Juan, M.C. & Alcaniz, M. (2005). ParSys: a new particle system for the introduction of on-line physical behaviour to three-dimensional synthetic objects. *Comput. Graphics—UK* 29 pp. 135–144.

Ritter, E.M.; Kindelan, T.W.; Michael, C.; Pimentel, E.A. & Bowyer, M.W. (2007). Concurrent validity of augmented reality metrics applied to the fundamentals of laparoscopic surgery (FLS). *Surgical endoscopy*, 21, 8, pp. 1441-1445.

Ro, C.Y.; Toumpoulis, I.K.; Ashton, R.C.Jr.; Jebara, T.; Schulman, C.; Todd, G.J.; Derose, J.J. Jr. and McGinty J.J. (2005). The LapSim: a learning environment for both experts and novices. *Stud. Health Technol. Inform.*, 111, pp. 414-417.

Rosen, J.; Solazzo, M.; Hannaford, B. & Sinanan, M. (2002). Biomedical Paper Task Decomposition of Laparoscopic Surgery for Objective Evaluation of Surgical Residents' Learning Curve Using Hidden Markov Model. *Computer Aided Surgery*, 7, pp. 49-61.

Rosser, J.C.; Rosser, L.E. & Savalgi, R.S. (1997). Skill acquisition and assessment for laparoscopic surgery. *Archives of Surgery*, 132, 2, pp. 200.

Sánchez-Peralta, L.F.; Sánchez-Margallo, F.M.; Moyano-Cuevas, J.L.; Pagador, J.B.; Enciso-Sanz, S.; Sánchez-González, P.; Gómez-Aguilera, E.J. & Usón-Gargallo, J. (2010). Construct and face validity of SINERGIA laparoscopic virtual reality simulator. *Int JComput Assist Radiol Surg*; 5, 4, pp. 307-15.

Satava, R.M.; Cuschieri, A.; Hamdorf, J. (2003). Metrics for objective assessment: Preliminary summary of the Surgical Skills Workshop. *Surgical endoscopy*, 17, 2, pp. 220-226.

Satava, R.M. (2008). Historical Review of Surgical Simulation—A Personal Perspective. *World journal of surgery*, 32, 2, pp. 141-148.

Schijven M & Jakimowicz J. (2003). Construct validity - Experts and novices performing on the Xitact LS500 laparoscopy simulator. *Surgical Endoscopy and Other Interventional Techniques*, 17, 1, pp. 803-810.

Schreuder, H.W.; van Dongen, K.W.; Roeleveld, S.J.; Schijven, M.P. & Broeders, I.A. (2009). Face and construct validity of virtual reality simulation of laparoscopic gynecologic surgery. *Am JObstet Gynecol*. May,200,5, pp. 540.e1-8.

Scott, D.J.; Bergen, P.C.; Rege, R.V.; Laycock, R.; Tesfay, S.T.; Valentine, R.J.; Euhus, D.M.; Jeyarajah, D.R.; Thompson, W.M. & Jones, D.B. (2000). Laparoscopic training on bench models: better and more cost effective than operating room experience? *Journal of the American College of Surgeons*, 191, 3, pp. 272-283.

Seymour, N. E.; Gallagher, A. G.; Roman, S. A.; O'Brien, M. K.; Bansal, V. K.; Andersen, D. K. & Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Annals of Surgery, 236, 4, pp. 458-463.

Sherman, V.; Feldman, L.S.; Stanbridge, D.; Kazmi, R. & Fried G.M. (2005). Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. *Surgical Endoscopy*, 19, 5, pp. 678-682.

Sidhu, R.; Grober, E.; Musselman, L. & Reznick, R. (2004). Assessing competency in surgery: where to begin? *Surgery*, 135, 1, pp. 6-20.

Sinigaglia, S.; Megali, G.; Tonet, O.; Pietrabissa, A. & Dario, P. (2005). Defining metrics for objective evaluation of surgical performances in laparoscopic training. *International Congress Series 2005*, Elsevier, pp. 509-514.

Sokollik, C.; Gross, J. & Buess, G. (2004). New model for skills assessment and training progress in minimally invasive surgery. *Surgical endoscopy*, 18, 3, pp. 495-500.

Stone, R. & McCloy, R. (2004). Ergonomics in medicine and surgery, *BMJ*, 328, 7448, pp. 1115-1118.

Stylopoulos, N.; Cotin, S.; Maithel, S.; Ottensmeyer, M.; Jackson, P.; Bardsley, R.; Neumann, P.; Rattner, D. & Dawson, S. (2004). Computer-enhanced laparoscopic training system (CELTS): bridging the gap. *Surgical endoscopy*, 18, 5, pp. 782-789.

Tavakol, M.; Mohagheghi, M.A. & Dennick, R. (2008). Assessing the skills of surgical residents using simulation. *Journal of surgical education*, 65, 2, pp. 77-83, 1931-7204, United States.

Tendick, F.; Downes, M.; Goktekin, T.; Cavusoglu, M. C.; Feygin, D.; Wu, X.; Eyal, R.; Hegarty, M. & Way, L. W. (2000). A Virtual Environment Testbed for Training Laparoscopic Surgical Skills. *Presence*, 9, 3, pp. 236-255.

Teschner, M.; Kimmerle, S.; Heidelberger, B.; Zachmann, G.; Raghupathi, L.; Fuhrmann, A.; Cani, M. P.; Faure, F.; Magnenat-Thalmann, N.; Strasser, W. & Volino, P. (2005). Collision detection for deformable objects. *Computer Graphics Forum*, 24, 1, pp.61-81.

Thijssen, A.S. & Schijven, M.P. (2010). Contemporary virtual reality laparoscopy simulators: quicksand or solid grounds for assessing surgical trainees?, The American Journal of Surgery, 199, 4, pp. 529-541.

Vassiliou, M.C.; Feldman, L.S.; Andrew, C.G.; Bergman, S.; Leffondré, K.; Stanbridge, D. & Fried, G.M. (2005). A global assessment tool for evaluation of intraoperative laparoscopic skills. *The American Journal of Surgery*, 190, 1,, pp. 107-113.

Verdaasdonk, E.G.G.; Stassen, L.P.S.; Monteny, L.J. & Dankelman, J. (2006). Validation of a new basic virtual reality simulator for training of basic endoscopic skills. *Surgical endoscopy*, 20, 3, pp. 511-518.

Verdaasdonk, E.G.G.; Stassen, L.P.S.; Schijven, M.P. & Dankelman, J. (2007). Construct validity and assessment of the learning curve for the SIMENDO endoscopic simulator. Surgical endoscopy, 21, 8, pp. 1406-1412.

Woodrum, D.T.; Andreatta, P.B.; Yellamanchilli, R.K.; Feryus, L.; Gauger, P.G. & Minter R.M. (2006). Construct validity of the LapSim laparoscopic surgical simulator. *Am. J Surg.*, 191, 1, pp. 28-32.

Youngblood, P.L.; Srivastava, S.; Curet, M.; Heinrichs, W.L.R.; Dev, P. & Wren, S.M. (2005). Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. *Journal of the American College of Surgeons*, 200, 4, pp. 546-551.

Zhang, A.; Hünerbein, M.; Dai, Y.; Schlag, P.M. & Beller, S. (2008). Construct validity testing of a laparoscopic surgery simulator (Lap Mentor): evaluation of surgical skill with a virtual laparoscopic training simulator. *Surgical endoscopy*, 22, 6, pp. 1440-1444.

Virtual Reality
Edited by Prof. Jae-Jin Kim

ISBN 978-953-307-518-1
Hard cover, 684 pages
Publisher InTech
Published online 08, December, 2010
Published in print edition December, 2010

Technological advancement in graphics and other human motion tracking hardware has promoted pushing "virtual reality" closer to "reality" and thus usage of virtual reality has been extended to various fields. The most typical fields for the application of virtual reality are medicine and engineering. The reviews in this book describe the latest virtual reality-related knowledge in these two fields such as: advanced human-computer interaction and virtual reality technologies, evaluation tools for cognition and behavior, medical and surgical treatment, neuroscience and neuro-rehabilitation, assistant tools for overcoming mental illnesses, educational and industrial uses In addition, the considerations for virtual worlds in human society are discussed. This book will serve as a state-of-the-art resource for researchers who are interested in developing a beneficial technology for human society.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ignacio Oropesa, Pablo Lamata, Patricia Sánchez-González, José B. Pagador, María E. García, Francisco M. Sánchez-Margallo and Enrique J. Gómez (2010). Virtual Reality Simulators for Objective Evaluation on Laparoscopic Surgery: Current Trends and Benefits., Virtual Reality, Prof. Jae-Jin Kim (Ed.), ISBN: 978-953-307-518-1, InTech, Available from: http://www.intechopen.com/books/virtual-reality/virtual-reality-simulators-for-objective-evaluation-on-laparoscopic-surgery-current-trends-and-benef

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.