Article

JDMS

Journal of Defense Modeling and Simulation: Applications, Methodology, Technology
2015, Vol. 12(3) 319–334

DOI: 10.1177/1548512914546148
dms.sagepub.com
SAGE

A Generic Architecture and Validation Considerations for Tactical Combat Casualty Care Serious Games

Hwa Feron[1], Axel Lehmann[1] and Florent Josse[2]

Abstract

Enthusiasm for high-tech, simulation-based training should never obscure the fact that trainees must be validly and dependably trained to perform tasks to given standard metrics. The educational performance and effectiveness of any such training method, simulation model and training scenario must therefore be rigorously verified and validated. Stringent quantitative acceptance criteria should be applied to allow fair and detailed skill-training effectiveness comparisons with traditional training methods. The challenges of verifying and validating a proposed generic architecture for serious games for effective virtual Tactical Combat Casualty Care (TCCC) are explored with regard to established tactical, medical and educational validation practices. As the fitness of such games for their educational purposes can make the difference between life and death of both the trainee and the casualty on the battlefield, their development and validation processes cannot be left to chance. Yet conflicting validation practices of medical, educational and military authorities may obscure design constraints for serious games combining medical and tactical purposes. Despite this uncertainty, our perception of best practice for establishing acceptance criteria for such games is outlined. Based on a minimum core validation procedure and official TCCC performance metrics, it emphasizes designing physiologically valid casualty models of deliberately limited complexity for TCCC training purposes, interacting with highly realistic serious game tactical training scenarios.

Keywords

TCCC, serious game, medical training simulation, verification and validation

1. Introduction and motivation

Effective battlefield emergency medical and tactical training according to the principles of Tactical Combat Casualty Care (TCCC)[1–4] requires teaching every soldier basic first responder lifesaving skills for safely stabilizing, directly on the battlefield, just a few lethal injury types demanding immediate care, such as hemorrhage, airway obstruction and pneumothorax. This approach has saved many lives in recent conflicts but tends to stress established military medical training structures because of the high costs and resource requirements needed to train the entire force, not just combat medical personnel.

As training interactions with a virtual patient are cheaper, safer and easier to implement and duplicate, the development of tactical and medical skill improvement games (surveyed, e.g. in our previous work[5] and in Graafland et al.[6]) has been driven by a demand for affordable yet realistic interactive training of large numbers of trainees. Such games potentially reduce requirements for training personnel, training grounds or expensive specialized equipment. Patient and soldier safety concerns as

[1]Institut für Technische Informatik, Universität der Bundeswehr München, Germany
[2]Sektion Notfallmedizin, Abteilung für Anästhesiologie und Intensivmedizin, Bundeswehrkrankenhaus Ulm, Germany

Corresponding author:
Hwa Feron, Institut für Technische Informatik, Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, Neubiberg, 85577, Germany.
Email: hwa.feron@unibw.de

well as a relative lack of live patients and active battlefields for training the required skills have contributed to this evolution.

1.1. Quality assurance

Incorrect, substandard or insufficient TCCC medical training would cost the casualty's life, while incorrect TCCC tactical training would endanger the trainee and his team in a combat situation. Therefore the necessity of rigorous educational quality control of games, physiological models and scenarios is internationally acknowledged, e.g. by the North Atlantic Treaty Organization (NATO) working group on advanced training technologies for medical healthcare professionals (NATO HFM-215):[7] 'all these issues require extensive attention to educational design principles, … [and] to rigorous validation to ensure that platforms are both safe and efficacious'.

A 'rigorous validation' of TCCC training games certainly would be a worthwhile goal, if only it were clearly defined, standardized and routinely practiced. Graafland et al.[6] observed in their systematic review of 30 serious games for medical education and surgical skills training that 'none of them had completed a full validation process for the purpose of use'. In fact, such a validation depends on the acceptance of several intersecting professional communities (military, medical, educational and simulation), each with its own practices and tradition, and each aware of its own priorities and liabilities, thus causing uncertainty and confusion as to the requirements to be met. For instance, concerning the validation of a TCCC manikin for the US army, Anton et al.[8] observed: 'No consensus has been reached as to the appropriate methodologies to provide test and evaluation strategies for verification, validation and accreditation.' The latest AMSP-01(B),[9] explicitly aiming to be the main modeling and simulation standards guideline document for NATO allies, highlights the necessity 'to align national and international efforts on V&V [verification and validation]; cultural differences of nations are slowing down the elaboration of international standards'. A primary motivation for this work has therefore been the development of a set of verification and validation principles for TCCC serious games which attempts to conform to established best practice and NATO validation guidelines in this field, as perceived by our team.

1.2. Training effectiveness dependence on pedagogical goals

Several large-scale evaluation and validation efforts by Sotomayor et al.[10,11] have revealed that TCCC serious games, when used solely as practical exercise to support traditionally taught TCCC topics, succeed in increasing motivation and reducing training time (e.g. for TCCC Simulation (T3CSim)).[12] They furthermore suggested benefits if used as refresher course (e.g. Trauma Connect (TraumaCon)).[13] Importantly, although those games were shown not to have contributed to any additional knowledge acquisition efficiency, these efforts also confirm the lack of negative effects on learning, thus apparently making the tested methods (T3CSim, TraumaCon and Virtual Medical Simulation Training Center (VMSTC)[14]) ethically acceptable for large-scale *supplementary* training, thus encouraging further development of TCCC training games.

Attempts to *replace* traditional or manikin-based instruction in whole or in part by wholly computer-based methods have however proved less successful in terms of training effectiveness. For instance, the effectiveness of VMSTC as *alternative* training was reported by Sotomayor et al.[11] to be significant, but not quite as good as conventional methods. Serious game use as alternative training may well be less costly, but would be more ethically problematic than its use as mere support for traditional training if its hidden costs were less perfect training and lost lives.

It is not yet clear exactly which factors limit the training effectiveness of serious games when used as standalone TCCC teaching methods, but a goal of the present study is to propose solutions and strategies grounded on published third-party work in an effort to better meet the ethical and practical TCCC validation and teaching effectiveness requirements for such uses.

By more systematically defining and formalizing core validation challenges and constraints specific to TCCC, it appears possible to develop suitable architectures methods, metrics and guidance leading to a clear roadmap for the design, verification and ethically correct validation of TCCC serious games, including better adapted metrics and measures of effectiveness. This roadmap is intended to help us to identify and overcome the current weaknesses of our attempt at a TCCC physiological model for a 'train as you fight' TCCC serious game demonstrator[5] for the German Federal Armed Forces.

2. TCCC training goals and principles

In a typical TCCC training game scenario, a small group including trainees (e.g. in a convoy or on patrol in hostile terrain), possibly accompanied by non-player characters (NPCs), e.g. medical evacuation team, hostile forces and civilian bystanders, comes under attack from small arms fire, improvised explosive devices (IED), rocket-propelled grenades (RPG), etc. and some members of the group sustain injuries.

Trainees are expected to apply as a team their tactical knowledge to regain control of the situation in the game so as to be able to safely approach the casualties and

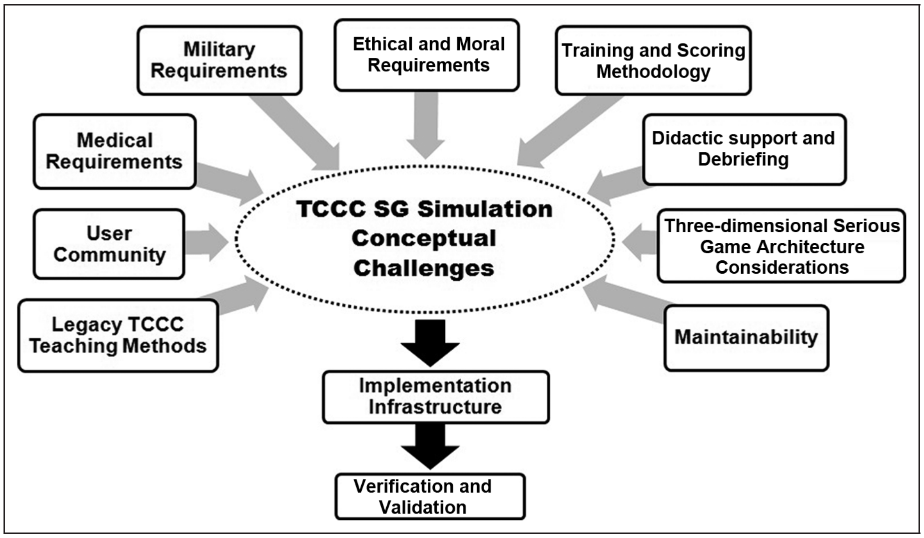

Figure 1. TCCC serious game conceptual challenges.

provide medical care according to TCCC. Proper 'train as you fight' TCCC training thus requires a serious game capable of extremely realistic battle situation simulation, with appropriate weapons, vehicles and NPC behavior modeling, especially intelligent enemy behavior modeling.

Having restored a modicum of safety, the trainees are expected to use their triage and emergency medical knowledge to diagnose, treat and save as many casualties as possible until arrival of the medical evacuation team. The game must therefore additionally be capable of sufficiently realistic casualty injury modeling, casualty consciousness modeling and casualty physiological modeling, with appropriate user interfaces to permit rapid diagnosis, triage and emergency battlefield care. Game vital signs fidelity and the reproduction of their sometimes confusing human variability are especially critical. Trainees may otherwise inadvertently train themselves to recognize game vital signs patterns not commonly found in real life, risking confusion and ineffectiveness when confronted with more ambiguous real vital signs. Life-threatening diagnosis errors and wrong or delayed emergency treatment would be a consequence of merely 'good enough', imperfect or overly repetitive vital signs simulations. Each completed scenario must be followed by an exhaustive debriefing.

3. TCCC serious game design and validation constraints

An evaluation of the tactical and medical training performance prerequisites of TCCC training leads to a series of game requirements, concrete limitations, verification issues and simulation and validation challenges conditioning the design and implementation of a TCCC serious game, as shown in Figure 1.

Addressing these challenges and issues in more detail below, we try to define a suitable quantitative validation metric for TCCC serious games which will orient our implementation. This leads in section 4 to a TCCC serious game generic architecture. Verification and validation procedures suitable for this application are then discussed in section 5, along with limits on the confidence that the training resulting from this game can be trusted by all parties involved.

3.1. Game scope limitations and the necessity of hybrid training

Ideally, a self-sufficient TCCC serious game would not only teach all TCCC procedures, but also test the trainee's overall TCCC proficiency once the curriculum is completed. However, attempting to teach all aspects of TCCC on a simulator might not be feasible or desirable. The civilian medical establishment normally requires some degree of actual practice on actual patients rather than a mere demonstration of learning on a simulator, in order to recognize full medical proficiency. Indeed, the last stage of Miller's pyramid[15] suggests that real-life medical experience, not simulation, should be the last step before examination and trainee certification.

For similar reasons, training resulting from the full-scale rehearsal of military maneuvers (field exercises) remains essential for maintaining preparedness, despite higher costs than virtual simulation techniques and war-gaming. Likewise, military medical training authorities[16] realize that the teaching effectiveness of serious game virtual world simulation heavily varies with the specific TCCC skill being taught and the capabilities and limitations of the human–machine interface inherent to the game hardware under consideration. They therefore advocate hybrid TCCC training, supplementing serious game training with at least some hands-on, real-life practice on manikins or part trainers, necessary for example for training bandaging and dexterity skills, and some full-scale field exercises for properly training ‘care under fire’, casualty transport and evacuation. The scope of TCCC serious games should therefore be restricted to what they do best, so as to better fit within a wider TCCC curriculum, since they are just one tool in an array of alternative and truly complementary training methods.

3.2. Compatibility with legacy TCCC training standards, methods and examination metrics

Regardless which training method (traditional or electronic) is used, trainees, be they soldiers with enhanced medical skills or medical service personnel like paramedics and military physicians, must be reliably trained, examined and ultimately certified to perform tasks to a given official standard. Examples of such standards include TCCC guidelines[2,3] and TC8-800 (medical education and demonstration of individual competence (MEDIC))[17] for US TCCC trainees, Tactical Rescue and Emergency Medicine Association (TREMA) guidelines[4] and Weisung für die Allgemeine Grundausbildung im Sanitätsdienst der Bundeswehr (AGASanDstBw)[18] for German forces.

In particular, TC8-800[17] describes an examination and competence scoring method for TCCC trainee medical skills validation testing for TCCC trainee certification, and asserts that ‘trainee skills validation also validates the training program to which they have been subjected’. This pragmatic statement was successfully applied in a large-scale survey by Phrampus et al.[19] which showed US army medics to perform to ‘acceptable clinical performance standards’ on TCCC lifesaving tasks, thus validating the manikin-based training methods which were in use at the time.

In that survey, the TCCC exam scoring sheets appear to have been interpreted to constitute a recognized validation metric for training effectiveness comparison of the medical aspects of TCCC legacy methods, such as manikin-based training. A similar TCCC test and validation metric (Figure 2), if generalized to cover all further aspects of TCCC (tactical, legal, ethical, etc.) in similar form, as sequences of essential steps for trainee examination, could form the backbone of a more rigorous quantitative validation and comparison method suited for TCCC serious games. It furthermore has the potential to be slightly adapted by medical military training experts and authorities to the specific norms, customs and exams of the country of application.

In Germany, the medical service of the armed forces (Sanitätsdienst der Bundeswehr) defines[18] military physician training standards, but organizations like TREMA, the Red Cross, and especially the German society for defense medicine and defense pharmacy—association of the German medical officers (Deutsche Gesellschaft für Wehrmedizin und Wehrpharmazie (DGWMP–VdSO)) also possess practical expertise in all aspects of battlefield emergency medicine. The TCCC training provided by serious games to German forces must in addition take into account other important training standards, such as the central service regulations of the German Federal Armed Forces (Zentrale Dienstvorschrift (zDV)),[20] which were intended to describe the sum of knowledge expected of German servicemen (medical, tactical, etc.); further detailed surgical procedure standards are available from the German trauma society (Deutsche Gesellschaft für Unfallchirurgie (DGU)).[21]

3.3. User community (customization of user needs)

TCCC trainees can be expected to exhibit very significant prior medical and tactical training and experience differences,[22,23] because TCCC is intended to be taught to all military service members: ordinary soldiers receive only 40–60 hours of TCCC training, but medical responders such as paramedics and physicians belonging to the medical corps (Sanitätsdienst) may have received up to eight years or more of formal medical education. The validation of the teaching emphasis and the learning objectives therefore must depend on trainee origin and user category.

As noted by Pettitt et al.,[24] casualty physiological model complexity has a price as well as a validation cost, which also depends on the level of medical proficiency to be trained. Considering that the largest number of trainees is trained to lower proficiency levels, a deliberate restriction of TCCC serious games to the trainee categories of lower medical proficiency might improve cost effectiveness while still serving the needs of a majority of trainees.

Finally, trainers, commanders, as well as medical and tactical experts require specialized interfaces for customizing scenarios or for in-game observation of the proficiency of the trainees.

3.4. Medical requirements

According to the TCCC <C>–A–B–C–D–E algorithm,[3,4] vital signs are an essential part of the basis on which any combat life saver has to make his triage and

Decompress the Chest: Needle Decompression Scoring Sheet					
Task		Completed			Interface
*Critical Elements		P	F	Score P/F	Input / Output
1.	*Took/verbalized body substance isolation (BSI) precautions.			5/0	I: Glove selection
2.	*Assessed the casualty to ensure the progressive respiratory distress was due to a penetrating chest wound.			5/0	I: Examine casualty (vital signs and wounds front/back) O: Avatar chest wound
3.	* Identified the second intercostal space (ICS) on the anterior chest wall at the mid-clavicular line (MCL) on the same side as the injury, approximately two finger widths below the clavicle.			5/0	I: Point to Avatar ICS zone
4.	Cleansed the site with an antimicrobial solution.			1/0	I: Clean wound
5.	Inserted the needle into the chest.			1/0	I: Needle selection and insertion in ICS zone O: Pop sound, hiss of air
6.	* Stabilized the catheter hub to the chest wall with adhesive tape.			5/0	I: Needle removal and adhesive tape positioning
7.	Placed the casualty in a sitting position or on their injured side (recovery position) during transport.			1/0	I: Avatar set in recovery position
8.	Removed their gloves and disposed of them appropriately.			1/0	I: Glove removal
9.	Documented the procedure on the appropriate medical form.			1/0	I: Input proper data on form
10.	* Did not cause further injury to the casualty			5/0	I: Survey wrong inputs

Figure 2. Needle decompression scoring sheet adapted from TC8-800.[17]

treatment decisions. The symptoms, time of death, reactions and vital signs of each casualty in a TCCC game should be governed by the evolution of its modeled vital parameters, as managed by a *casualty physiology model* appropriately reacting to each TCCC-relevant injury type and realistically responding to treatment actions, e.g. tourniquet application, clearing airway, etc.

Such a TCCC casualty physiology model would include at least three interlinked sub-models: a hemorrhage model driven by blood loss and overall casualty blood volume, an oxyhaemoglobin saturation model able to simulate lung function, pneumothorax and airway obstruction effects on blood oxygen concentration and a crude shock and consciousness model driven by the previous two models. These sub-models attempt to plausibly and realistically model the evolution of the rapidly evolving vital parameters (e.g. blood oxygenation; overall blood volume; consciousness states such as alert, responsive to verbal stimuli, responsive to painful stimuli and unresponsive (AVPU); shock; etc.) for each injury and each casualty, in a trauma-dependent, time-dependent and treatment-dependent manner. At each update cycle, each casualty medical model generates from the current status of these vital parameters plausible observable vital signs (heart pulse, breathing noises, lip color, reactions to stimuli, etc.) to allow proper diagnosis, avoiding misleading artifacts. Slower evolving conditions, such as hypothermia/exposure can be simply preprogrammed/scripted.

The casualty medical model should focus on proper modeling of the TCCC-relevant injuries, on correct or at least plausible output of those vital signs necessary to their proper diagnosis and on realistic modeling and interfacing of treatment actions and their consequences. Quite a few of these models are already available; their detailed implementation[25] and the criteria for their verification and validation[26–28] are the subject of ongoing research which is out of the scope of the present work.

Simulated casualties are expected to display amputations, wounds, the effects of those wounds on consciousness, speech, noises and vital signs, as well as the effects of treatments chosen from medical aid materials which must be added to game assets.

To this end, each standard TCCC procedure[2,3] to be trained must be examined in great detail, based for instance on the detailed TCCC procedure scoring templates,[17] in order to determine a medical content validation/medical Measure of Performance (MoP) table[8] specifying exactly which casualty simulation capabilities are required in terms of visual, tactile and acoustic symptoms, vital signs and specific input/output interfaces.

For instance, Figure 2 presents a possible quantitative examination metric of tension pneumothorax procedure steps, along with the human–machine interface needs for each treatment step, as derived from a standard 'tension pneumothorax' scoring template from TC8-800,[17] DA Form 7595-23-R, for STP 8-68W13-SM-TG, Task 081-833-0049. This metric merely differs from Phrampus et al.[19] by the higher weight assigned to 'critical elements', i.e. essential procedural steps, the minimum passing total score being dependent on the number of such steps. In addition, the last column of Figure 2 specifies the required capabilities of the human–machine interface for proper tension pneumothorax diagnosis and scoring at each step, for building a list of Measures of Performance (MoP) for content validation purposes.

However, as noted by Pettitt et al.,[24] attempting to model more than what is strictly required is not only costly, it may cause unneeded validation problems and expense, and may, as noted by Sotomayor et al.,[29] end up consuming too much processing power.

Manikins for TCCC training, such as CAESAR[30] and SimMan3G Mystic,[31] have also been designed to implement most of these required visual, tactile and acoustic simulation capabilities.

To determine the medical content validation list, at least the following steps must be carried out in order to define the essential medical simulation capabilities needed for game verification:

- lists of steps to be scored (e.g. according to Figure 2) for each relevant TCCC medical procedure (both diagnosis and treatment);
- list of essential vital signs and symptoms to be simulated (e.g. as suggested by the last column of Figure 2), as derived from the above lists;
- list of standard diagnosis aids available to military staff;
- medical materials and material interfaces to be used for each medical procedure;
- corresponding instruction sets for each medical procedure and trainee category; and
- items and capabilities required by the different scenarios in the curriculum.

3.5. Military requirements

A major training advantage of virtual scenarios over manikins is to alert the trainee to specific battlefield conditions without danger or expense by providing a set of assorted scenarios combining typical medical and tactical training challenges suitable to train the whole range of TCCC know-how, including returning fire in the 'care under fire' phase. For instance, Helm et al.[22] propose a series of just four approved scenarios: combat situation with facial gunshot wound causing upper airway obstruction, IED attack with severed limbs, RPG attack with massive bleeding and shrapnel injuries and traffic accident involving three wounded soldiers. The tactical verification and validation requirements derive from the 'train as you fight' need for TCCC training to take place under stressful battlefield conditions, such as winter nights or under (simulated) enemy fire. The practice of performing TCCC test procedures[17,19] on manikins in full battle dress realistically prepares trainees to rapidly deal with clothing impediments and other real-world obstacles to proper palpation, diagnosis of vital signs and wound identification.

3.6. Ethical and moral requirements

Medical responders with battlefield experience, legal experts and chaplains should be considered for face validation of scenarios preparing trainees to make tough life and death choices and to face legal or moral dilemmas[18] on the battlefield which may gravely affect TCCC decisions. These may include triage duties with multiple severely wounded casualties, some of which must die, mission or tactical imperatives conflicting with emergency care needs, etc.

3.7. Psychological and pedagogical requirements (training and scoring methodology)

Simulation-based serious games are based on a large ecosystem of interacting models creating a believable and realistic virtual world capable of providing a suitable environment for training.

Of course, the training scenarios and the federation of models resulting in the simulation of the battlefield for TCCC game purposes must be subjected to a standardized validation process, e.g. Simulation Interoperability Standards Organization (SISO) Generic Methodology for Verification and Validation (GM–VV).[9] However, even very general, emerging simulation validation methodologies such as GM–VV do not seem to be optimally suited to validate certain aspects of serious games, because a serious game is in fact much more than a mere simulation creating a suitable game environment:

- Serious games must include motivating and educative elements of game-play, and must offer a success/failure or win/lose incentive, or else they are not games, but mere simulations with reduced training and motivation value.

- Serious games aim to be training methods and must therefore also be validated by following the standards[6] established for validating training methods and their components.

Proper TCCC training requires the game to provide not only an appropriate and self-coherent educational sequence of training scenarios teaching the entire spectrum of tactical and medical skills, it must also provide a suitable scoring method[19] rewarding the correct and timely sequence[17] of emergency care steps. If TCCC scoring were to reward the wrong actions (e.g. rapid death of the casualty), the game would still be teaching the wrong message despite possibly using entirely valid and reliable physiological models. Therefore, the validation of a serious game cannot be reduced to the mere validation of the simulations on which it is based, and must in addition be validated according to established procedures for validating training methods. The scoring method must reward in a pedagogical and validated manner the proper application of the principles embodied in TCCC and must encourage proper behavior of the trainee, including incentives for early and correct completion of the triage tasks. The scoring rules must therefore take into account the evolution of the simulated vital signs, shock, etc. and must monitor the trainee reactions to them as a function of elapsed time.[19]

3.8. Didactic support/debriefing

A proper debriefing is considered essential[22] for TCCC scenarios. Debriefing validation is straightforward in the case of standardized feedback (such as a video explaining how an expert would have reacted). However the validation challenges become significant when performance evaluation results are personalized in a trainee-dependent manner, by an expert system taking into account the mistakes, previous experience and previous training records of the trainee. The ability to replay critical parts of the trainees' actions is nevertheless expected to permit constructive feedback and discussion of the mistakes that may have been made.

3.9. Three-dimensional serious game architecture: requirements, interoperability, reuse and standards

The complexity of serious games for combined medical and tactical training purposes is significantly greater (artificial intelligence (AI), hundreds of pieces of equipment and character models) than that of games for surgery training purposes, because the latter can dispense with modeling the environment outside of the patient's body. TCCC training games may require modeling an entire three-dimensional Afghan city environment on a busy market day with realistically animated functional objects such as vehicles, weapons and accurate weapon dynamics, as well as realistic-acting civilian or hostile NPCs. Indeed, they must simulate the tactical aspects of TCCC which are essential in, and inextricably linked to, the TCCC 'care under fire' phase, which may include a gun battle. If designed from scratch, often at great expense, the highly interdependent models and game modules must not only be individually verified, validated and documented; they also must be shown to interact correctly with one another.

Such tactical modeling capabilities in fact already exist in three-dimensional, first-person shooter multiplayer serious games for tactical training (e.g. the Virtual Battle Space 2 (VBS2)-based serious games running on the Simulation and Test Environment of the Bundeswehr (SuTBw))[32] which can be adapted and leveraged for TCCC purposes (e.g. VBS2 2.0[33] includes an amputation simulation system, wound textures, as well as a medical kit for restoring fellow soldiers to full health, a tourniquet for stopping bleeding, etc.). These existing game packages document procedures for adding content and come with large libraries of various validated game features, characters, effects, terrain, controls and equipment. They would be advantageously completed by a few additional in-game TCCC user interface modules for interacting with a casualty physiological model simulating vital signs, bleeding, impaired breathing and various levels of consciousness in all game casualties, as well as additional resources from simulation libraries (e.g. from the NATO Simulation Resource Library (SRL)).[34]

In addition, serious games for tactical training typically require players to be familiar with a very complex user interface. Reusing such tactical serious games also for TCCC training in addition to tactical training therefore not only simplifies the validation of the game interface, but also saves trainees the otherwise substantial time spent relearning yet another one: it permits teaching TCCC in the ergonomically familiar virtual environment from the trainee's previous tactical training.

Our proposed TCCC serious game architecture (Figure 3) supports a multiplayer tactical game because training TCCC 'care under fire' operations usually requires both teamwork (perimeter security and caring for casualties) and the full tactical/weapons/enemy NPC capabilities of such games. However such multiplayer games often require specialized hardware, lengthy installations, high processing power and corresponding game servers requiring dedicated training facilities. Simplified TCCC games with lower resolution images and perhaps simplified scenarios and environments should be designed to run on leaner, portable trainee platforms so as to permit untethered hybrid training anywhere. Trainee software should be optimized for smaller

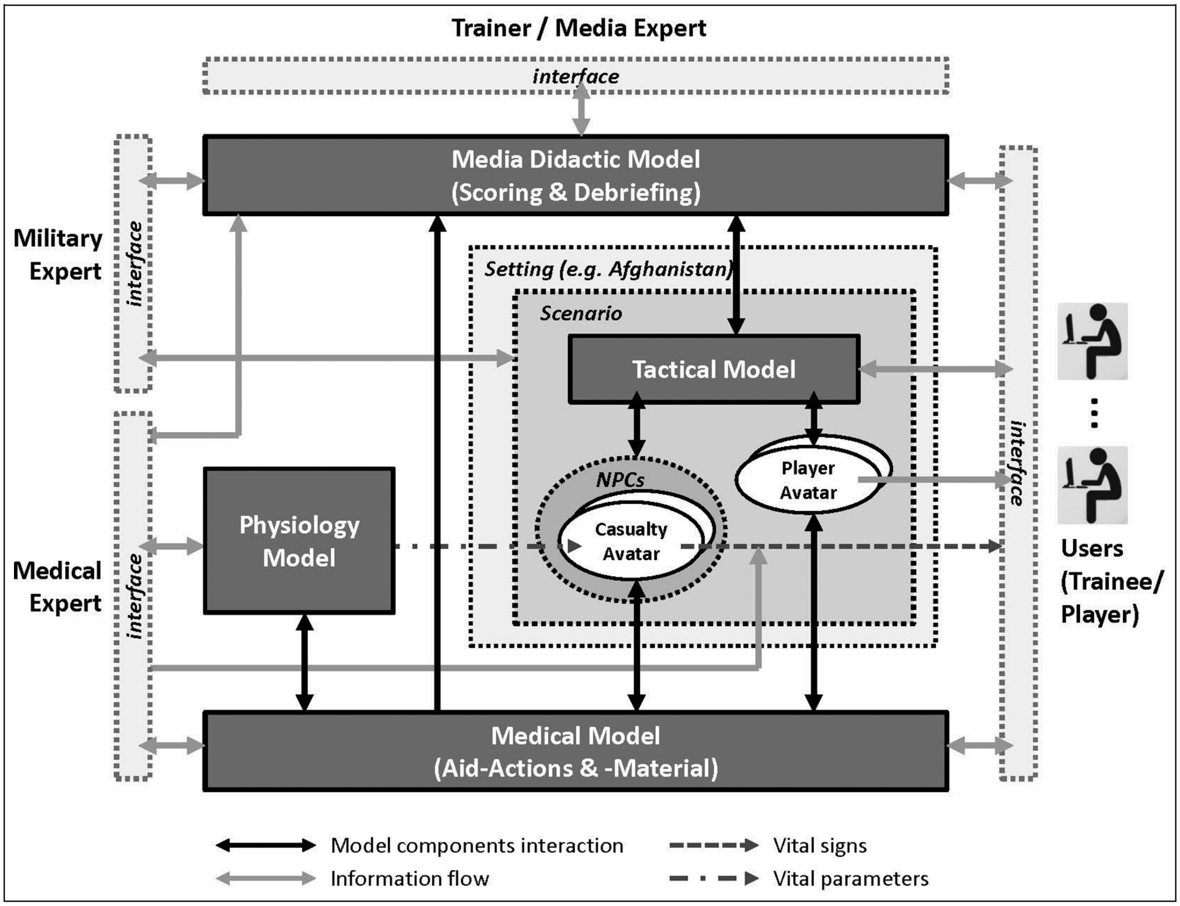

Figure 3. Generic architecture of our TCCC serious game simulation.

displays, weaker processors and portable implementation in multiplayer distributed architectures suitable for team training during field exercises, try to reuse validated components of existing games (AI, physics, assets, e.g. reuse NATO/Bundeswehr VBS2) and try to relocate most data and processing power to servers located, for example, on support vehicles. Unlike the primitive point-and-click interfaces of traditional serious game hardware, the built-in sensors (e.g. accelerometer, compass, Global Positioning System (GPS), camera and touch screen) and actuators (e.g. tactile feedback of vibration alarms for more realistic heart beat simulations) of such modern serious game platforms as smart phones or wearable soldier systems offer opportunities for more elaborate and user friendly human–machine interfaces which may be leveraged to better simulate vital signs and facilitate interactions with the casualty.

The reuse and adaptation of an existing multiplayer tactical serious game for TCCC simulation and training also implies the reuse of all its interoperability standards, in particular interoperability with databases storing virtual representation of real-world terrain matching that of the current field exercise, or with flight simulators for medical evacuation helicopters.

For maximum interoperability, portability, updateability and reuse while minimizing validation needs, standard interfaces (ASMP-01(B): NATO modeling and simulation standards profile) and interoperable data standards, e.g. Synthetic Environment Data Representation and Interchange Specification (SEDRIS)[35] should be applied.

3.10. Maintainability

TCCC instructions are frequently updated, the details of the medical procedures they teach change as well, along with the performance evaluation method; further, the range of skills being taught is expanded at every TCCC publication. The methods used to teach the procedures are even likely to change more frequently. Therefore both the TCCC game and its validation methodology must be highly modular and clearly documented, e.g. based on the Institut für Technik Intelligenter Systeme (ITIS) Leitfaden (guideline for model documentation)[36] and designed so as to be easily extendable and modifiable, as well as must include the corresponding scenarios, introductions and debriefing materials. The typical villains change from conflict to conflict, and so do the cultural environments of typical serious games.

4. A TCCC serious game example: the generic architecture

Our TCCC serious game generic architecture (Figure 3), designed to meet the above-defined challenges and validation constraints, consists of the following elements:

- Tactical Model;
- Physiology Model;
- Medical Model;
- Media Didactic Model;
- 4 Interfaces.

A **Tactical Model**, driven by a scenario. It controls the essential tactical terrain game objects, character avatars and environment aspects in the three-dimensional virtual world, e.g. hostile forces, equipment, physics and game effects; it is advantageously implemented by reuse of an existing three-dimensional multiplayer game (e.g. VBS2, Emergency[37]), which admits the next three elements as game modifications (mods).

A (patho-) **Physiology Model** for each casualty. It realistically simulates the evolution of vital parameters and other physiological effects of the TCCC injuries and generates the physiological effects of the treatments and erroneous treatments on the evolution of the vital functions. Since the plausibility and correctness of the generated data is crucial to proper diagnosis, stringent validation of this simulation against experimental data would ideally be needed. However, obtaining actual human vital data on life-threatening injuries and their effects on human consciousness is ethically difficult or impossible due to the dangers to the test subjects. This can instead be achieved, as already suggested by Hester et al.,[28] by comparing the simulations outputs with the outputs of an already validated integrative physiology model such as HumMod, or by face validation through medical experts.

A **Medical Model**, which receives injury information for each casualty from the scenario, treatment action data from the trainee and vital parameters from the physiology model. It is in charge of managing, interpreting and converting this information, thus controlling the appearance, vital signs, movements, reactions and sounds of each casualty avatar in a manner consistent with its injuries, pain level, applied treatments, consciousness and physiological status as output by the separate physiological model. In particular, it uses this information to control injury appearance and location on the casualty avatar, to compute the effects of the TCCC-relevant trauma, as well as to compute the effectiveness of any aid action for modification of the relevant vital parameters of the physiology model, such as changes in bleeding rate.

A **Media Didactic Model**, is in charge of scoring and debriefing: it assesses the diagnosis and treatment decisions of the trainee by comparing them to the normal sequence of TCCC actions, grades his performance and provides feedback, especially on errors and omissions in the sequence of treatment steps.

In addition, four kinds of **interfaces** are foreseen, for the trainees to interact with the game, for medical and military experts to perform validation and updates (e.g. changed TCCC rules, changed military equipment or strategies and model adjustments to better reflect reality) and for trainers, pedagogy and media experts. This last interface is intended for updating the debriefing materials such as training videos for changing training incentives by modifying scoring parameters or for influencing the course of a multiplayer game, e.g. adding or removing enemies in certain positions.

Due to the necessary reuse of existing tactical serious games and their standard user interface, the capabilities of the human–machine interface (HMI) clearly will be limiting the range of TCCC skills trainable by means of serious games. Despite the more versatile interfaces made possible by smart phone technologies, this fundamental limitation continues to preclude proper training and automatic assessment of dexterity skills like bandaging. It is therefore one of the main reasons why TCCC serious games can only be considered as one component of a hybrid training curriculum, and not as a standalone, self-sufficient training solution capable of teaching and scoring the whole spectrum of TCCC skills.

However, the architecture of Figure 3 was designed to be HMI-independent, because it is meant to be compatible with standard point-and-click workstations and with devices with richer interfaces such as smart phones. Specifically, the hybrid training philosophy requires TCCC topics taught by means of serious games to be restricted according to the inherent hardware interface limitations of each game system and results of concurrent validation testing, and thus may lead to different training scopes and contents on different training platforms. Specifically, for each given system, the concurrent validation tests of Figure 4 with the scoring metric of Figure 2 provide a way to decide which TCCC actions should not be taught by means of serious games, whenever traditional methods yield significantly better training effectiveness, while the list of input/output requirements in the last column of Figure 2 provides clues as to whether a given HMI has the required minimum capabilities to permit training of a given TCCC skill.

Validation Topics (suggested by references)	Aspects	Evaluation Panel/Data Source	QAL/QTY Evaluation
Are the teaching scenarios useful for TCCC?[38]	Test for each scenario, eliminate unneeded features[24,29]	Domain experts	Qualitative (QAL)
Is the game teaching all relevant aspects of TCCC? *Content validity* [6, 8, 38]	Consider all scenarios: Any omitted skills? Are they suitable for serious games training?	Medical experts, tactical experts	Y/N
Is the game properly teaching every aspect of TCCC?[8,38]	Test each aspect e.g. triage, injury types, etc.	Trainees results, exam scores	Quantitative (QTY)
Has the player learned what he was supposed to?[8,38] *Predictive validity surrogate*	Is the trainee learning wrong, odd or unintended behaviors?[10,11,19,41]	Trainees results (e.g. from exams)	QTY
Is the performance evaluation/ scoring appropriate? *Construct validity*[6,38]	Test each scoring aspect and interactions, compare scores of trainees and experts[19]	Trainees results, domain experts	QTY
Learning from mistakes, compulsory debriefing[22]	Pedagogical/didactical debriefing effectiveness	Trainees results, re-test	QTY
Is the game sufficiently motivating and user friendly? *Face validity* (gaming aspects)[6]	Gaming ergonomy, entertainment and motivation value[10,11]	Trainees (questions and answers)	QAL
Is the game better than traditional TCCC training? *Concurrent validity*[7,38]	Test and compare each aspect, may be relativized by cost matters[24]	Trainees results comparison with legacy methods	QTY
Medical realism, fidelity.[8, 22, 38] *Medical external validity*[24] *as Predictive validity surrogate*	Vital signs plausibility, injuries, consciousness, aid-actions and -materials, etc.	Medical experts opinion, exam results	QAL
Tactical realism and fidelity: 'train as you fight'[22, 38] *Tactical external validity as predictive validity surrogate*	Tactical aspects of TCCC: combat stress, team security, IED, scenarios, NPC behavior, equipment, terrain, etc.	Tactical experts opinion	QAL
Ethical and legal issues, moral dilemmas[18] *Ethical/legal face validity as predictive validity surrogate*	Life and death aspects built-in or undesired in scenarios, triage, legal obligations, basis for decisions, etc.	Ethical experts opinion, trainee reactions	QAL
Coherence of the scenarios with non-TCCC skills, team or tactical training and military rules *Common sense face validity as predictive validity surrogate*	Is the game inadvertently conveying or promoting bad practices in other fields or skills, or harming teamwork ?[10,11,41]	Tactical and other experts from neighboring fields	QAL

Figure 4. TCCC serious game validation considerations.

5. TCCC serious game verification and validation

Since proper training is a matter of life and death for both casualties and first responders, it must be ensured that trainees are properly learning all required skills. The game must also be proven to actively discourage incorrect behavior, such as tactically unsafe actions during the TCCC 'care under fire' phase, and must be proven to systematically provide plausible, typical and sufficiently realistic simulated patterns of vital signs which are suitable for diagnosis.

5.1. Validation metrics and expectations

TCCC training games components (scenarios, scoring methods, didactic materials and models) must of course be individually subjected to a stringent verification and quality assurance early in the design of the game. They also must be proven to be compatible with each other: their combination must be validated as training software and demonstrated to be at least as correct and effective in its training results as traditional training methods in the following domains:

- emergency medical competences (e.g. detailed first aid procedures);
- TCCC tactical competences (safety in the 'care under fire' phase and IED awareness); and
- ethical/legal/emotional competences (triage and tactical/medical life and death dilemmas, e.g. 'care under fire', no-win scenario and mission imperatives over Hippocratic oath).

Trainee competences acquired through game training must be compared[10,11,19] (e.g. by subjecting trainees to comprehensive exams[17]) to those attained by a traditionally trained control group, such as a group trained with manikins, in order to be able to validate all aspects of the training method as an effective training method at least for a well-defined subset of TCCC skills.

A TCCC training method relying on scenarios with simulated casualties in a simulated hostile environment must in addition be able to sufficiently faithfully represent real-world scenarios, enemies, characteristic TCCC injuries and situations, in order for the trainee not to feel disoriented, disturbed or powerless when similar situations occur in the real world.

5.2. Training validation methodologies

For establishing the validity of serious games intended for medical training purposes, a consensus emerged in the medical serious game literature around a jointly developed method[6,39] by the American Psychological Association (APA), the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME). It seems to be well accepted worldwide, particularly for validation of surgical training, e.g. for laparoscopy psychomotor training skills,[38] also in the military: NATO and the US Department of Defense Defense Modeling and Simulation Office (DMSO) use it for verification, validation and accreditation purposes (chapter 10.2 of RTO-TR-HFM-128)[40] of military medical simulations. Anton et al.[8] describe a similar US army validation approach on a manikin for TCCC training, first checking whether the method is teaching everything it is supposed to (by defining Measures of Performance) and then measuring how effective it is at doing so (by defining Measures of Effectiveness). Other assessment methods[10,11] have been successfully used for evaluating the training effectiveness of TCCC serious games, but the APA method, being endorsed both by NATO for verification, validation and accreditation purposes and by the world medical community for the evaluation of medical education and surgical skill training, remains our preferred choice.

5.3. Specific verification and validation criteria for TCCC serious games

At a minimum, the following validations should be carried out for different TCCC scenario situations and different injuries, as insistently recommended by, for example, Graafland et al.[6] and Verdaasdonk et al.[38] by means of:

- *a rigorous content validation*, to *verify* that the game includes all prerequisites enabling it to meet its TCCC design objectives, in terms of features, in particular modeling and scoring features. For instance, section 3.4 details how the medical content validation list, which permits definition of the relevant Measures of Performance,[8] can be derived from validated TCCC examination templates such as those of TC8-800,[17] as explained in the practical example of Figure 2;
- *a face validation* strategy which includes a validation of the vital signs plausibility for different injuries as well as for normal conditions;
- *a construct validation* strategy to demonstrate whether the game performance scoring scheme (Figure 2) assorted with the model capabilities permits differentiating between novices and experts;
- a meaningful *concurrent validation* strategy based on a TCCC-relevant metric,[17–19] (e.g. Figure 2). The average final score of trainees subjected to different training methods is expected to supply a

rational basis for deciding whether a TCCC game-based method is better and more efficient than a traditional TCCC teaching method or not. Metric-based concurrent validation with respect to validated manikin-based training methods and traditional training methods determines whether the game is acceptable in whole or in part, which parts of it (training topics, e.g. tension pneumothorax, interface aspects, etc.) must be improved and which parts are not fit for use and must be discarded;

- some approximation of *predictive validation*: although acquired medical competence can be tested[19] on manikins in battle dress, the overall acquired competence usually cannot be tested in real battlefield conditions on real human casualties. *Surrogate* predictive validation tests, e.g. on unintended effects,[19,41] realism[38] and believability[42] seem needed, as discussed below.

In particular, the qualitative and quantitative evaluations listed in Figure 4 should be carried out by well-chosen panels in order to ensure the game can be trusted.

5.4. Individual game component validation during game development

As the training validation methodologies imply large-scale, costly and lengthy trials as well as a fully implemented final version of the game to test on trainees, they cannot be of any help to guide the game designers during implementation.

Other more practical and economical quality assurance approaches are therefore needed to control the development efforts, which would otherwise be blind. In particular, strategies for validation of the individual components of the game (scoring, scenarios, physiology model, tactical simulation, virtual reality representations, user interface, etc.) will have to be continuously performed during the development process, long before the final training method validation, in order to ensure the viability of game development progress. The necessary effort can be mitigated by maximizing reuse of already developed and validated game components.

Crucially, game designers need early information on the validity of the physiology simulation components which simulate the effects of injuries and treatments on casualties. As the trainees are supposed to reach their TCCC treatment decisions based on casualty vital signs, particular attention must be devoted to defining measurements and limits or concrete boundaries, preferably quantitative wherever feasible, for the quality (plausibility) and pedagogical suitability (proximity to a textbook case evolution) of the simulated vital signs trajectories and other simulated injury symptoms. Obviously, an invalid casualty vital signs simulation would cause the players to acquire serious misconceptions and might cause them to become confused or reach incorrect diagnoses in real-world situations, thereby endangering the casualty's life.

These quantitative quality assurance criteria imply some economical way of determining and modeling vital sign trajectories and symptoms having textbook reference character for each TCCC injury category and severity.[27–29]

A major validation challenge remains wherever the access to reliable experimental calibration data is restricted to only part of the domain of applicability. For instance, even the most advanced pathophysiological hemorrhage simulations[43] have not been, and cannot be, experimentally validated for unconscious casualties due to ethical concerns, yet models simulating hemorrhage in unconscious patients are needed for TCCC. In those cases, pragmatic validation solutions have to be identified for the remainder of the TCCC domain of applicability to provide surrogate reference data for comparison and validation. This may include, for example, data provided by an integrative model trusted by the medical community even outside of its calibration range, where no trustworthy or ethically acceptable experimental data is available.

In particular three approaches could be combined as follows:

- observation/measurement: trying to reuse experimental case data or wartime records to test plausibility;
- expert opinion: face validation feedback from training experts and battle experienced military physicians and medical responders; and
- model comparison: compare outputs[29] with one or more reference physiological models.

This combination of approaches is essential to offset and mitigate their individual limitations:

Observation/measurement data often turns out to be anecdotal or incomplete due to an uncontrolled environment (e.g. battlefield), fragmentary records, unspecified context, statistically insufficient numbers of clinical cases or ethical limitations, e.g. on life-threatening medical experimentation on humans, thus limiting its usefulness.

Face validation may yield confusing results as different experts may have different views and dissimilar clinical, first aid and battlefield experience and levels of expertise.

Reference physiological models trusted by the medical community may prove less reliable in really life-threatening vital signs ranges due to the ethical impossibility of performing suitable validation of their outputs in such ranges in controlled environments. Rigorous quantitative validation remains unattainable but probably can be satisfactorily approximated.

5.5. Predictive validity measurement difficulties

'The proof of the pudding is in the eating;' i.e. the only real test of something is as what it is intended to be used for. Although acquired medical competences can be quantitatively and qualitatively tested[19] on advanced manikins, many limitations and advantages of any TCCC teaching method are likely to show up only on real battlefields in the treatment of non-standard injuries on actual human casualties.

Predictive validity testing is essential[6] to measure the degree of concordance of a game outcome/examination score with task performance in reality. Although it is easily measured[41] for serious games teaching vehicle handling, for TCCC serious games it remains mostly inaccessible. TCCC game outcome certainly can be quantitatively measured[19] based on a validated scoring system derived, for example, from the step-by-step evaluation sheets in TC8-800,[17] while the player is not in the heat and stress of actual combat. Yet the value, actual effectiveness and limitations of any TCCC teaching method are likely to show up only on real battlefields, as relatively infrequent and non-reproducible instances of actual lifesaving performance under fire. This important measure of TCCC training effectiveness in general is not reliably measurable, because assessment of training effectiveness and trained skills of the trainee in general cannot be carried out on actual human patients. Factors preventing such assessment in a true battlefield environment include personnel safety, lack of casualties with wounds suitable for standardized TCCC testing, statistically insufficient data, insufficient control of experimental conditions and the impracticality of conducting experiments or gathering their results during an actual battle.

It follows that the APA validation method[6,39] cannot be fully applicable for validation of TCCC training because the true predictive validity measurement is not easily accessible. This is why Figure 4 does not include any actual predictive validation and why it attempts to approximate the purpose of predictive validation by proposing alternative heuristic surrogate measures.

5.6. External validity, realism and believability

Simulation realism is often considered to be very important[8,22,38] but may be limited by processing power[29] on portable hardware. It might be considered to be not only relevant to face validation, but also as a surrogate measure of predictive validity, on the premise that higher levels of realism of the interaction of models creating the game world in the image of the real world would have to influence any measure of similarity between game outcomes and real outcomes.

Indeed, improper trainee preparation to battlefield conditions due to insufficient simulation realism can be dangerous. It can negatively affect predictive validity by causing undue stress and disorientation of the trained combat lifesaver in real-life combat situations, thus endangering the casualty (insufficiently mastered skills), the trainee himself (insufficient preparedness and inadequate precautions or cover from enemy fire) or the rest of the team (by attracting enemy fire). The freshly trained lifesaver may feel hindered by the stress and constraints of real combat, surprised by the unanticipated weight and reactions of the casualty, disturbed by the amount of blood flow from real wounds and challenged to identify vital signs in the suboptimal environment of a noisy battlefield at night if these items had not been adequately simulated in the game.

In a serious game, these aspects also relate to the qualitative concepts of external validity, simulation realism, simulation fidelity or believability and depend in particular on the interaction of a very large number of models in order to provide a believable representation of reality. These concepts have to be considered as essential for certain aspects of battlefield effectiveness of serious game training, but are hardly quantifiable, except in terms of required processing power.

Feinstein and Cannon[42] argue in their conclusion that simplistic simulations with low fidelity may on the contrary assist novices 'by focusing their attention on important variables'. Conversely, this can be interpreted to suggest (see also Pettitt[24]) that simplistic simulations may be less suitable for TCCC training games training non-novices such as military physicians.

Moreover, undue additional stress and confusion may be caused by discovering conditions on the battlefield which were not properly simulated by a simplistic simulation and which may cause precious seconds to be wasted in time-critical situations, e.g. while restoring the casualty's airway. Finally, the most 'important variable' for the trainee practicing TCCC is tactical rather than medical, because tactical 'variables' related to carelessness are more likely to kill him, e.g. an enemy bomb or bullet. For this reason, the TCCC serious game really should be based not on an overly simplistic simulation but on a sophisticated tactical serious game with clever enemies armed with realistic weapons, in order not to convey a false sense of security to the trainees during the 'care under fire' simulation, which might cost them their life on an actual battlefield. Thus a compromise between trainee terminal portability and realism requirements must be found.

6. Conclusion

TCCC serious games need to accurately model complex tactical and medical life-threatening situations and scenarios for training purposes. We propose a generic architecture suitable for TCCC serious games which combines TCCC-specific pathophysiological modeling solutions with off-the-shelf tactical game components. We further explore validation challenges for TCCC serious games and indicate how to arrive at a list of game scoring and verification and validation criteria, including surrogate predictive validation criteria, by showing how official examining strategies,[17,18] for measuring the proficiency of a TCCC trainee which have been defined and validated for traditional teaching methods in great step-by-step detail, can be reused and generalized for validating serious games and for comparing their performance with those methods.

Statistical comparisons of the trainees TCCC proficiency scores attained by serious games and traditional teaching methods are essential to promote acceptance of, and trust in, game-based training methods. These comparisons could help clarify for which injuries and skills TCCC immersive serious games might become more promising than traditional methods, thereby suggesting lower cost hybrid TCCC teaching curricula requiring less hands-on training.

However, further research work remains needed to determine how realistic and detailed the graphics and models really need to be for effective TCCC tactical and medical training, to better simulate the effects of injuries on consciousness and casualty reactions for diagnosis purposes and to define better surrogate measures of predictive validity.

Funding

This research is sponsored by the German Federal Armed Forces Project 1074: 'Simulationsbasiertes sanitätsdienstliches Training für Nicht-Sanitätspersonal und Sanitätspersonal' (Simulation-based medical training for non-medical personnel and medical personnel). The views expressed in this paper are those of the authors. They do not necessarily reflect the official policies or positions of the German Federal Armed Forces (Bundeswehr) or NATO.

References

1. Butler FK. Tactical combat casualty care: evolving concepts and battlefield experience. *Military Med* 2007; 172(11): 1.
2. US Department of Defense. Tactical combat casualty care guidelines. *Military health system*. http://www.naemt.org/Libraries/PHTLS%20TCCC/TCCC%20Guidelines%20120917.sflb (2012, accessed 5 March 2014).
3. PHTLS: Prehospital trauma life support. Military 7th ed. NAEMT, Mosby/JEMS, 2010.
4. Tactical Rescue and Emergency Medicine Association (TREMA). Richtlinien der TREMA e.V. für TCCC (Taktischen Verwundetenversorgung). V1.2. http://46.38.238.62/wp-content/uploads/2013/06/TREMA-e.V.Guidelines-fuer-TCCC-1.2.pdf (2012, accessed 5 March 2014).
5. Feron H and Hofmann M. Tactical combat casualty care: strategic issues of a serious simulation game development. In: *2012* Winter Simulation Conference, Berlin, Germany, 9–12 December 2012.
6. Graafland M, Schraagen JM and Schijven MP. Systematic review of serious games for medical education and surgical skills training. *Brit J Surg* 2012; 99(10): 1322–1330.
7. Hodges J. Human factors and medical panel meeting on advanced training technologies for medical healthcare professionals—research technology group 215 (NATO-RTO HFM-RTG-215), workshop proceedings, London, UK, 2012.
8. Anton JJ, Burns E, Norfleet J, et al. US army research, development, and engineering command (RDECOM) independent test and evaluation of the stand alone patient simulator (SAPS) under the DoD challenge program. In: *2009 ITEA live-virtual-constructive conference: requirements for modeling and simulation in test and evaluation*, El Paso, TX, USA, 12–15 January 2009.
9. NATO Standardization Agency. AMSP-01 NATO modelling and simulation standards profile. Edition (B), version 1, NATO Allied Modelling and Simulation Publication, 2012.
10. Sotomayor TM and Proctor D. Assessing combat medic knowledge and transfer effects resulting from alternative training treatments. *J Def Mod Sim Appl Meth Tech* 2009; 6(3): 121–134.
11. Sotomayor TM, Salva AM and York BW. Measuring the training effectiveness of combat lifesaver simulation training systems. In: *2012 interservice/industry training, simulation, and education conference (I/ITSEC)*, Orlando, FL, USA, 3–6 December 2012.
12. Engineering and Computer Simulations (ECS). Tactical combat casualty care simulation (TC3Sim). http://www.ecsorl.com/solutions/tactical-combat-casualty-care-simulation (2010, accessed 5 March 2014).
13. Engineering and Computer Simulations (ECS). Trauma connect (TraumaCon). http://www.ecsorl.com/solutions/trauma-con-combat-lifesaver-mma (2012, accessed 5 March 2014).
14. Engineering and Computer Simulations (ECS). Virtual medical simulation training center (VMSTC). http://www.ecsorl.com/solutions/virtual-medical-simulation-training-center-vmstc (2010, accessed 5 March 2014).
15. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990; 65(9): 63–67.
16. Hoitz J. Simulation in military medicine. International military forum. http://www.mci-forum.com/category/experiences/419-Simulation_in_military_medicine_more_than_just_skill_training.html (2012, accessed 5 March 2014).
17. Headquarters, US Department of the Army. Medical education and demonstration of individual competence (MEDIC), Training circular (TC) No. 8-800, 2009.
18. Ocker K. Weisung für die Allgemeine Grundausbildung im Sanitätsdienst der Bundeswehr (AGASanDstBw) mit integrierter Einsatzvorbereitender Basisausbildung für Konfliktverhütung und Krisenbewältigung und die Ausbildung Sanitätsgrundlagen. Inspekteur des Sanitätsdienstes der

Bundeswehr, Bundesministerium der Verteidigung (German Federal Ministry of Defence), 2006.
19. Phrampus P. Ft. Sam 91 Whiskey combat medic medical simulation training quantitative integration enhancement program, Technical US Army Medical Research and Materiel Command, Fort Detrick, MD, 2011.
20. Bundesminister für Verteidigung Inspektion des Sanitäts- und Gesundheitswesens II. Verbandlehre, Zentrale Dienstvor schrift 49/23, German Federal Ministry of Defence, 1960.
21. Deutsche Gesellschaft für Unfallchirurgie. S3-Leitlinie Polytrauma/Schwerverletzten-Behandlung (guideline on treatment of patients with severe and multiple injuries), AWMF Register-Nr. 012/019, 2011.
22. Helm M, Lührs J, Josse F, et al. Konzept zur Basisausbildung von Notärzten im Sanitätsdienst der Bundeswehr. *Notfall + Rettungsmedizin* 2012; 15: 146–151.
23. Inspekteur des Sanitätsdienstes der Bundeswehr. Weisung Ersthelferausbildung von Nicht-Sanitätspersonal (draft), Bundesministerium der Verteidigung (German Federal Ministry of Defence), 2009.
24. Pettitt BH, Norfleet J and Descheneaux CR. Task specific simulations for medical training: fidelity requirements compared with levels of care. In: *Interservice/industry training, simulation, and education conference (I/ITSEC)*, Orlando, FL, USA, 30 November–3 December 2009.
25. Feron H. Casualty vital signs modeling for serious game-based battlefield medical training simulation. In: *Zwischenbericht zur Studie: Simulationsbasiertes sanitätsdienstliches Training für Nicht-Sanitätspersonal und Sanitätspersonal*. Munich, Germany, April 2012.
26. Feron H, Lehmann A and Hofmann M. Challenges of and criteria for validating a physiology model within a TCCC serious game. In: *2013* Winter Simulation Conference, Washington, DC, 8–11 December 2013.
27. Hester RL, Iliescu R, Summers R, et al. Systems biology and integrative physiological modeling. *J Physio* 2011; 589(5): 1053–1060.
28. Hester RL, Brown A, Husband L, et al. HumMod: a modeling environment for the simulation of integrative human physiology. *Front Physio* 2011; 2: 12.
29. Sotomayor TM, Quintero B, Salva AM, et al. High fidelity physiological model for immersive simulation and training. In: *2012 interservice/industry training, simulation, and education conference (I/ITSEC)*, Orlando, FL, USA, 3–6 December 2012.
30. CAE Healthcare. CAESAR: point-of-Injury trauma simulator. http://www.medsimlab.com/brochuras/caeser.pdf (2011, accessed 5 March 2014).
31. Laerdal Medical. SimMan3G mystic. http://www.laerdal.com/us/SimMan3GMystic?q=SimMan%C2%AE%203G%20Mystic (2013, accessed 5 March 2014).
32. Hauxen H and Janssen H. Simulation and test environment of the Bundeswehr (SuTBw). Transforming defence through modelling and simulation—opportunities and challenges (STO-MP-MSG-094), NATO Science and Technology Organization, 2012.
33. Bohemia Interactive Australia Pty Ltd. Virtual battlespace 2. http://manuals.bisimulations.com/vbs2/2-00/devref/#Adding_Models/AM_Adding_Content.htm (2012, accessed 5 March 2014).
34. NATO. Recommendations on the establishment of a NATO simulation resource library (SRL), RTO TR-051, AC/323(MSG-012)TP/04, NATO RTO, 2003.
35. ISO/IEC JTC 1. Synthetic environment data representation and interchange specification (SEDRIS), ISO/IEC18023, 2006.
36. Lehmann A, et al. Leitfaden für Modelldokumentation (guideline for model documentation), Abschlussbericht. – BMVg –Studienauftrag Nr. M/GSPO/2A024/2A924, SKZ 129902114, Institut für Technik Intelligenter Systeme GmbH, Munich, Germany, 2005.
37. Deep Silver (a division of Koch Media GmbH) and Serious Games Solutions GmbH. Serious game titled ‘‘emergency.’’ http://e2012.deepsilver.com/ (2014, accessed 5 March 2014).
38. Verdaasdonk EG, Stassen LP, Monteny LJ, et al. Validation of a new basic virtual reality simulator for training of basic endoscopic skills. *Surg Endosc Intervent Tech* 2006; 20(3): 511–518.
39. Gallagher AG, Ritter EM and Satava RM. Fundamental principles of validation and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc Intervent Techn* 2003; 17(10): 1525–1529.
40. NATO. Verification and validation. RTO technical report: human behavior simulation in constructive simulation (NATO RTO-TR-HFM-128), 2009.
41. Reber EA, Bernard B, McDowell P, et al. Improving naval shiphandling training through game based learning. In: *2012 interservice/industry training, simulation, and education conference (I/ITSEC)*, Orlando, FL, USA, 3–6 December 2012.
42. Feinstein AH and Cannon HM. Fidelity, verifiability, and validity of simulation: constructs for evaluation. *Dev Bus Sim Exper Learning* 2001; 28: 57–62.
43. Summers RL, Ward K, Witten T, et al. Validation of a computational platform for the analysis of the physiologic mechanisms of a human experimental model of hemorrhage. *Resuscitation* 2009; 80(12): 1405–1410.

Author biographies

Hwa Feron is a Korean-raised, French researcher at the University of the Federal Armed Forces Munich, Germany. After professional experience with health insurance databases at Blue Cross/Blue Shield of Western Pennsylvania and organ transplant databases at the Regional Organ and Tissue Bank, Rush-Presbyterian Hospital, Chicago, IL, she obtained a MS in AI at Illinois Institute of Technology, Chicago, IL.

Axel Lehmann, chair professor of modeling and simulation at the University of the Federal Armed Forces Munich, Germany until 2011, is executive director of Institut für Technik Intelligenter Systeme GmbH. He is a fellow of the German informatics society (Gesellschaft für Informatik e.V.).

Florent Josse is an anesthesiologist at the Federal Armed Forces Hospital Ulm (Bundeswehrkrankenhaus Ulm), Germany. He is specialized in emergency medicine and diving medicine. From 2005–2008, he was responsible for the training of combat first responders at the Special Operations Training Center at Pfullendorf. He is currently responsible for the TCCC training of medical service soldiers such as paramedics and emergency medical technicians (EMTs)–physicians at the military hospital in Ulm. He is a fellow of the TREMA and DGWMP–VdSO societies.