

## **CHALLENGES OF AND CRITERIA FOR VALIDATING A PHYSIOLOGY MODEL WITHIN A TCCC SERIOUS GAME**

Hwa Feron  
Axel Lehmann

Marko Hofmann

Institut für Technische Informatik  
Universität der Bundeswehr München  
Werner-Heisenberg-Weg 39, GERMANY

Institut für Technik Intelligenter Systeme  
Universität der Bundeswehr München  
Werner-Heisenberg-Weg 39, GERMANY

### **ABSTRACT**

Tactical Combat Casualty Care (TCCC) principles save lives on the battlefield but tend to stress established military medical training structures because of the need to train the entire force, thereby encouraging new large-scale computer-based training methods such as serious games. Since improper training would cause avoidable casualty deaths, important experimental efforts have been made to ensure the validity and reliability of these new methods and of their components. This survey of validation efforts attempts to identify best practice, challenges and limitations for proper design and rigorous validation of TCCC serious games and their virtual casualty pathophysiological simulation components, guided by the hippocratical requirement that new TCCC training methods must be more effective and reliable than traditional ones, in order for their large-scale use to be ethically acceptable (first, do no harm). TCCC serious game validation solutions are then deduced to guide our own TCCC serious game demonstrator design.

### **1 INTRODUCTION AND MOTIVATION**

The principles of Tactical Combat Casualty Care (TCCC) (Butler 2007, TCCC 2012, PHTLS 2010, TREMA 2012), essentially a series of simple life saving steps and clear priorities (triage, diagnosis and treatment) to be carried out in the minutes following life-threatening trauma, have greatly contributed to improving battlefield survival rates.

However they require appropriate TCCC training for nearly all servicemen, because medics cannot be present on every battlefield. This challenging goal of training the whole force is difficult to meet by traditional methods, i.e., with the help of trained instructors, not least because of budget limitations and of the relative scarcity of competent tutors. For instance, the German military medical authorities (“Sanitätssamt der Bundeswehr”) intend in a first phase to try to reduce tutoring needs by training at least part of the TCCC curriculum also by means of a serious game, without loss of training efficiency.

As a first step towards this goal, we previously reported (Feron and Hofmann 2012) on our development strategy for a battlefield first aid training game for the German Federal Armed Forces, relying on separate development teams working in parallel, a pedagogical expert team concentrating on deciding how and in which form the medical principles are to be taught, and a game developer team best able to package that subject-matter in an attractive game with a motivating storyboard and an appealing graphics environment.

This work explores some of the practical, technical and ethical challenges and necessities of its future validation, attempts to define some of the validation criteria which the serious game as well as its components (such as the scoring module and the casualty physiology simulation) must meet, and suggests architectures which might facilitate the application of these criteria.

## **1.1 TCCC Serious Game Validation Successes and Promises**

### **1.1.1 Serious Games Used as Supplementary Teaching in Addition to Standard Training**

Several large scale evaluation and validation efforts by Sotomayor and Proctor (2009) and Sotomayor, Salva and York (2012) have revealed that serious games such as T3CSim (2010), when used solely as practical exercise to support traditionally taught TCCC topics, succeed in increasing motivation and reducing total training time. They furthermore suggested that games such as TraumaCon (2012) are effective as TCCC refresher course. The experimental evaluations, based only on single casualty hemorrhage, airway management and pneumothorax first aid, appeared to imply that no additional knowledge acquisition was attributable to the tested games. Importantly, they also confirmed a lack of negative effects on learning, thus apparently making the tested methods (T3CSim, TraumaCon, VMSTC) ethically acceptable for large scale supplementary training, thus encouraging further development of TCCC training games.

### **1.1.2 Serious Games Used as Alternative to Standard Teaching for Certain TCCC Topics**

Attempts to replace traditional instruction methods in whole or in part by serious games have however proved less successful in terms of training effectiveness, but already seem to hold promises for the future, given the right pedagogical and model accuracy approaches. For instance, VMSTC (2010) was tested as an alternative to traditional methods by Sotomayor, Salva and York (2012): even though the measured training efficiency proved significant, it remained inferior to that achieved by traditional methods alone, thus making this kind of serious game use as alternative training method much more challenging and ethically problematic than its use as mere support for traditional training. It is not yet clear which factors limit the training effectiveness of serious games when used as standalone TCCC teaching methods, but at least the following causes of incorrect training require particular attention during game design and validation:

- Inadequate or erroneous vital signs and injury symptoms simulation.
- Unsatisfactory or insufficient medical interaction possibilities with virtual casualties.
- Improper, insufficient or absent feedback (TCCC scoring incentives or TCCC debriefing).
- Insufficient or inadequate didactical contents.

## **1.2 Ethical Requirements on Validation**

The above mentioned efforts confirm that serious game techniques have significant potential to efficiently teach, or help to teach, extremely complex subject matter, such as emergency medical care for a severely wounded, dying casualty, by enhancing game entertainment value to increase learning, training and discovery motivation. However, they also suggest pedagogical shortcomings in current game designs, which still suffer from important teaching efficiency limitations when used as standalone teaching methods.

Since substandard training can cost lives on a real battlefield, inadequate TCCC tactical or medical training would raise serious ethical, moral and liability concerns. Accordingly, the accreditation of TCCC training games cannot only rely on vague allegations of training effectiveness, qualitative face validation assertions or subjective arguments.

It is therefore essential for military medical authorities to gather rigorous, convincing and objective (if possible quantitative) rather than subjective validation data before training based on any particular TCCC serious game can be authorized, let alone generalized. How else could a mere game be trusted to teach behavior decisive upon life and death situations, including emergency surgery skills which in other settings are strictly reserved to medically trained personnel ?

For this reason, the validity, reliability and effectiveness of the TCCC serious game as training method and that of every one of its components must be backed up (within the limits of feasibility and within the ethical limits of medical experimentation) by convincing validation strategies, documenting at least the fitness of the training method as well as that of its components for their intended uses, along with their superiority over traditional and proven TCCC teaching methods.

Indeed a consensus is now developing among medical military e-learning users about the necessity of rigorous validation, as emphasized by the international medical panel in NATO project HFM-215 on Advanced Training Technologies for Medical Healthcare Professionals (Hodges 2012, 1): ***“all these issues require extensive attention to educational design principles,..[and] to rigorous validation to ensure that platforms are both safe and efficacious.”*** It is therefore the responsibility of the modeling and simulation community to try and make every effort to suggest the next best feasible validation solution for providing this kind of quality assurance, even if a perfect solution does not exist. This is what we attempt here.

## 2 VALIDATION OF SERIOUS GAMES AS TEACHING METHODS

Any serious game must offer a success/failure or win/lose incentive, or else it is not a game, it is a mere simulation with reduced teaching value. Thus, a serious game implementing such incentives is much more than a collection of models: it fundamentally aims to be a training method and must therefore be validated by following the standards established for validating training methods and their components.

Although validation strategies developed for simulations, models or model federations have to play a role, they are per se not sufficient because such a serious game must include motivating and educative elements of game-play, particularly a scoring method implementing in a pedagogical manner the knowledge contained in the TCCC guidelines. The TCCC scoring module therefore must be carefully validated as well: rewarding wrong actions, leading to a more rapid death of the casualty, would result in a game teaching the wrong lessons despite possibly using entirely valid and reliable physiology models.

The scoring method may be carried out by an appropriate expert system consisting of a set of scoring rules monitoring vital signs and trainee actions. The rules may, for instance, be derived from the grading sheets of TC8-800 (MEDIC 2009). They are expected to score trainee actions in dependence of simulated vital signs, shock and consciousness status, indications of severe blood loss, etc. The necessary validation of the scoring method or expert system must examine not only the individual rules but also their interactions. Such interactions occur in multiple casualty triage situations, or in casualties with multiple injuries.

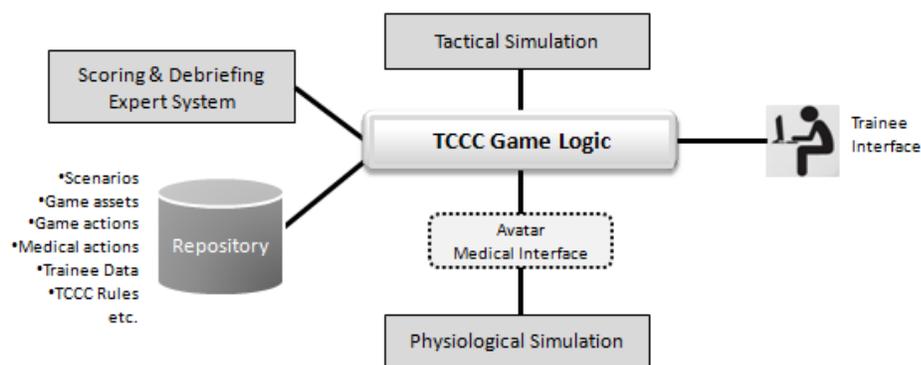


Figure 1: TCCC single player serious game structure.

A validation approach for training methods was developed over the past 40 years by the American Psychological Association (APA), the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME). It appears to have emerged as consensus in the medical serious game literature (Gallagher, Ritter, and Satava 2003; Graafland, Schraagen, and Schijven 2012), particularly for validation of surgical training, e.g., for training laparoscopy psychomotor skills

(Verdaasdonk et al. 2006). In defense applications, NATO and the US DoD DMSO also use this approach for VV&A of military medical training simulations (Chapter 10.2 of RTO-TR-HFM-128). Anton et al. (2009) 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 Training Performance) and measuring how effective it is at doing so (by defining Measures of Training Effectiveness).

For TCCC serious games validation, at least the evaluation procedures of Table 1 should be carried out for the different scenario situations and the injury categories relevant to TCCC training. They include:

- **a rigorous content validation**, to ensure the game includes all prerequisites enabling it to meet its TCCC design objectives, in terms of features, in particular modeling and scoring features;
- **a face validation** strategy which includes a validation of the vital signs **plausibilities** for different injuries as well as for normal conditions;
- **a construct validation** strategy to demonstrate whether the game performance scoring assorted with the model capabilities permit differentiating between novices and experts; and
- a meaningful **concurrent validation** based on a TCCC-relevant metric, capable of deciding whether a TCCC game-based method is better and more efficient than a traditional teaching method or not. Concurrent validation determines whether it is acceptable in whole or in part, which parts of it must be improved, and which parts are not fit for use and must be discarded.

Table 1: Evaluation procedures for TCCC serious game validation.

Validation Topics	Aspects	Evaluation Panel/ Eval. Data Source	QAL/QTY Evaluation
Are the teaching scenarios useful for TCCC ?	Test for each scenario	Domain experts	QAL
Is the game teaching all aspects of TCCC?	Consider all scenarios, <i>Content validity</i>	Medical experts, Tactical experts	Y/N
Is the game properly teaching every aspect of TCCC ?	Test each aspect (triage, exhaustive list of injury types)	Trainees results & exam scores	QTY
Has the player learned what he was supposed to ?	Is the trainee learning wrong, odd or unintended behaviors ?	Trainees results (e.g., from exams)	QTY
Is the performance evaluation/scoring appropriate ?	Test each scoring aspect, also interactions: <i>Construct validity</i>	Trainees results & domain experts	QTY
Learning from mistakes and debriefing	Pedagogical/didactical debriefing effectiveness	Trainees results & re-test	QTY
Is the game sufficiently motivating ?	Gaming, entertainment and motivation value, <i>Face validity</i>	Trainees (Q&A)	QAL
Is the game better than traditional TCCC training ?	Test and compare each aspect, <i>Concurrent validity</i>	Trainees results comparison	QTY
Medical realism	Vital signs, injuries, consciousness, aid actions & material, etc.	Medical experts opinion	QAL
Tactical realism	Tactical aspects: attacks, situations, missions, NPC and enemy behaviors, equipment, terrain, etc.	Tactical experts opinion	QAL
Ethical issues and moral dilemmas	Life and death aspects built-in or undesired in scenarios, triage, etc.	Ethical experts opinion	QAL
Coherence of the scenarios with non-TCCC skills and military rules	Is the game inadvertently conveying or promoting bad practices in other fields or skills ?	Tactical experts from neighboring fields	QAL

In detail, at least the qualitative and quantitative (QAL/QTY) evaluation procedures according to the aspects listed in Table 1 must be carried out by the appropriate panels or gathered from the appropriate evaluation data sources in order to build sufficient trust in the training method and identify its limitations.

### 3 INDIVIDUAL GAME COMPONENT VALIDATION DURING GAME DEVELOPMENT

The intended purpose of our serious game is to improve the life-saving performance of combat first responders by providing them with a motivating additional training possibility for the application of the TCCC algorithm. Since an *evaluation* (and training efficiency comparison with traditional training methods) of the serious game according to that purpose and as outlined above would have to assess performance measures of trainees before and after playing the game and their motivation to play it, these assessments can only be carried out after completion of game implementation, not during game design. Furthermore, these complex assessments require significant advance planning and large resources.

Other more practical and economical quality assurance approaches are therefore needed to guide game designers before game completion, or else the development efforts would have to proceed blindly. In particular, strategies for ad-hoc validation of the individual components of the game, shown in the simplified game structure example on Figure 1 (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 above described final training method validation, in order to ensure that game development progresses along a viable path. The necessary effort can be mitigated by maximizing game component reuse. In particular most elements of the TCCC tactical simulation component may already be available in other existing training games, for which they have been validated. Accordingly, it may be easier and more desirable (in order to offer combined training of tactical missions with a TCCC element) to incorporate TCCC training into an existing tactical training game such as VBS2 (2012).

Paramount among these other quality assurance approaches is the *validation* of the physiology *simulation components* of the serious game that 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, might cause them to become confused or to reach incorrect diagnoses in real world situations when observing significantly different real injury symptoms, thereby endangering the casualty's life.

These needed quantitative quality assurance criteria imply the necessity of identifying some economical way of determining and modeling vital sign trajectories and symptoms having textbook reference character for each TCCC injury category and severity. This can be done as described below.

### 4 VALIDATION OF TCCC COMPUTER SIMULATIONS

According to Schlesinger et al. (1979, 3), validation denotes the “*substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model*”. Validation is always driven by a specific domain of application, which is set by a purpose. Validating a model means to assess whether this purpose can be fulfilled or not.

The purpose and domain of applicability of the physiology model appears to be a simple and straightforward consequence of the intended purpose of the TCCC serious game, as follows:

- (1) The intended purpose of TCCC games is to teach proper battlefield diagnosis and treatment techniques for a few TCCC injuries to trainees having varying degrees of prior medical knowledge, based on a virtual casualty exhibiting simulated vital signs. Specifically, the purpose of the development of our serious game is to improve measures of lifesaving performance of combat first re-

- sponders (e.g., “Einsatzersthelfer B” in German military forces) on the battlefield in diagnosing and treating *only* the restricted set of TCCC injuries. Non-TCCC injuries are not to be treated.
- (2) Since the necessary manual skills are not realistically trainable via mouse or keyboard, the game must focus on the training of cognitive processes (injury diagnosis with vital signs evaluation, appropriate treatment decisions) associated with TCCC, not on aid action execution details. These *limitations on the domain of applicability* of the game suggest that a proper TCCC training curriculum involving serious games must be hybrid (Hoitz 2012), i.e., that it must also involve traditional training sessions, e.g., with manikins, to teach skills which are inherently unsuitable for computer-based training (bandaging skills, casualty manipulation and transport, etc.). Other good reasons for deliberately restricting the domain of applicability of the serious game could be caused by the non-availability of a validated scoring method or a proven inferior teaching efficiency of virtual simulations for a specific category of injury or treatment.
  - (3) The intended purpose of the physiological simulation is to support the TCCC teaching aims by generating parameters and symptoms which are *sufficiently typical* of those injuries to teach the trainee caring for a wounded virtual casualty how to identify whether or not the injury is a TCCC injury, to teach how to quickly and correctly diagnose which TCCC injury is being simulated, its degree of lethality for triage purposes, and which treatment options might be the correct ones.
  - (4) Consequently, the intended purpose of the physiology model is to provide the player with *realistic and sufficiently characteristic responses* (vital signs, injury appearance, casualty reactions) of a casualty according to different types of TCCC injuries and different TCCC aid actions within a domain of applicability to be defined clearly and precisely according to TCCC requirements (physiological modeling of vital parameters relative to a few TCCC injuries until death occurs).
  - (5) According to the Schlesinger definition, a “*satisfactory range of accuracy consistent with the intended application*” indicates that the original being modeled (in general, some part of reality) is always the crucial reference system whose processes have to be imitated with a purpose-specific degree of precision. This is a thorny challenge for the calibration of TCCC parameter modeling, because this reference may not be experimentally accessible in life-threatening parts of the domain of applicability, for instance due to ethical restrictions on human experimentation.
  - (6) Since access to reliable experimental calibration data may be restricted to only part of the domain of applicability, 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 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. For instance, HumMod hemorrhage simulations (Summers et al. 2009) have not been experimentally validated for unconscious casualties due to ethical concerns, but the TCCC game may have to simulate unconscious hemorrhaging casualties, using its output outside of its range.

#### 4.1 Performing Ad-hoc Physiology Validation with TCCC Surrogate Data in the Design Phase

Game designers need rapid and economical validation techniques to test their prototypes while designing them. In particular, the following three approaches, each insufficient on their own, could be combined:

- **Observation/Measurement:** reuse experimental case data or wartime records to test plausibility.
- **Expert opinion:** face validation feedback from training experts and battle experienced medics.
- **Model comparison:** compare outputs with one or more reference physiological models.

*Observation/Measurement data* is often anecdotal or incomplete due to an uncontrolled environment (battlefield), fragmentary records, unspecified context, statistically insufficient numbers of clinical cases, or ethical limitations on life-threatening medical experimentation on humans.

*Face validation* may prove confusing as different experts may have different views and experiences.

*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. Since rigorous quantitative validation remains unattainable at the extremes of the vital signs ranges, the three approaches must be combined to avoid errors.

#### **4.2 Validation of a Physiology Model Using a Trusted or Validated Reference Model**

Inherent complexity problems of integrative physiology modeling and validation may seem daunting but have in fact been the object of continuous and sustained experimental research.

One example is a growing family of related integrative physiology models developed mainly at the University of Mississippi, including the Guyton model (Guyton, Coleman, and Granger 1972), HUMAN model (Coleman and Randall 1983), Quantitative Circulatory Physiology (QCP) (Abram et al. 2007) and HumMod (Hester et al. 2011a). In particular, QCP internal hemorrhage and intravenous infusion simulations (critical also for TCCC training) have been used for years at the the University of Salamanca medical school (Rodriguez-Barbero and Lopez-Novoa 2008) for medical education purposes.

Specific validation criteria developed for the intended purpose of human integrative physiology modeling have also been suggested by that team of developers, as follows:

- Simulated vital signs must correspond qualitatively (expected range and direction of evolution) and quantitatively, both in steady state (unwounded or stabilized casualty) and in dynamics (over time during hemorrhage, pneumothorax evolution, etc.), to vital signs evolutions observed in experiments (Summers et al. 2009).
- Where validation with experimental data is impractical, model comparisons in accordance to the above principles with trusted reference models such as HumMod have been suggested (Brown et al. 2011).

Statistically, hemorrhage is the most deadly TCCC injury, making trustworthy validation of hemorrhage vital signs simulation an absolute priority. Although gaining experimental hemorrhage data from humans in controlled environments for model validation purposes might be unethical insofar as it would endanger human lives, surrogate experimental hemorrhage vital signs were more safely obtained from twenty 37 year old male volunteers by use of the Lower Body Negative Pressure (LBNP) technique (Summers et al. 2009). On this occasion, HumMod model outputs were shown to be within 5% of average experimental vital sign trajectories before the onset of unconsciousness, suggesting use of either the open source HumMod (or the closely related QCP trusted at the Salamanca medical school for hemorrhage training of several hundred students) as hemorrhage calibration reference and surrogate vital signs validation data source for TCCC physiology modeling, according to the procedure outlined in Hester et al. (2011b).

Other TCCC-relevant examples have been published, such as pneumothorax vital parameter evolution outputs produced by HumMod (Hester et al. 2011a), which might be considered to constitute plausible surrogate data for comparison with the pneumothorax outputs of our serious game demonstrator.

The interaction of HumMod with brain functionality may eventually become relevant for better modeling of TCCC consciousness symptoms such as AVPU (Alert/Verbal/Pain/Unresponsive) but has not yet led to validated applications, even though applications to military serious games are discussed as being the subject of ongoing research in Dancy, Ritter, and Berry (2012) and Hester et al. (2012)

The integration of HumMod with a 3D medical interactive simulation for teaching or medical serious game purposes has been described by Kofranek et al. (2011) in their “Atlas of physiology and pathophysiology” application. Their 3D simulator platform built around HumMod seems adaptable to TCCC games.

Crucially for our project, Sotomayor et al. (2012) present vital signs outputs of different models for the same upper arm amputation scenario, obtained by fitting the TCCC serious games (vMedic 2012, VMSTC 2010) with an interface to communicate with an external physiological model to enable a fast

comparison with the game's own vital parameter simulator, thus paving the way for a validation according to the principles suggested by Hester et al. (2011b). Their setup *has concurrent validation potential* to reveal vital parameter and vital signs deviations from trusted physiology models and suggest adjustments to the game model to reduce those deviations, without costly, lengthy and ethically questionable medical experimentation on humans.

### 4.3 Validation Architecture for the TCCC Serious Game Physiology Model

Using simpler, less accurate physiological models may sometimes be justified for TCCC games destined to run on platforms with limited processing power, such as mobile phones, but the vital signs outputs of those models must be compared to a trusted reference.

The casualty simulation structure of Figure 2 reapplies the external physiology model principle proposed by Sotomayor et al. (2012) together with the validation principles by comparison as developed by Hester et al. (2011b). It permits rapid and economical physiology model calibration and validation during the design phase, thus allowing the design team to test its concepts before a full-scale experimental validation can be carried out.

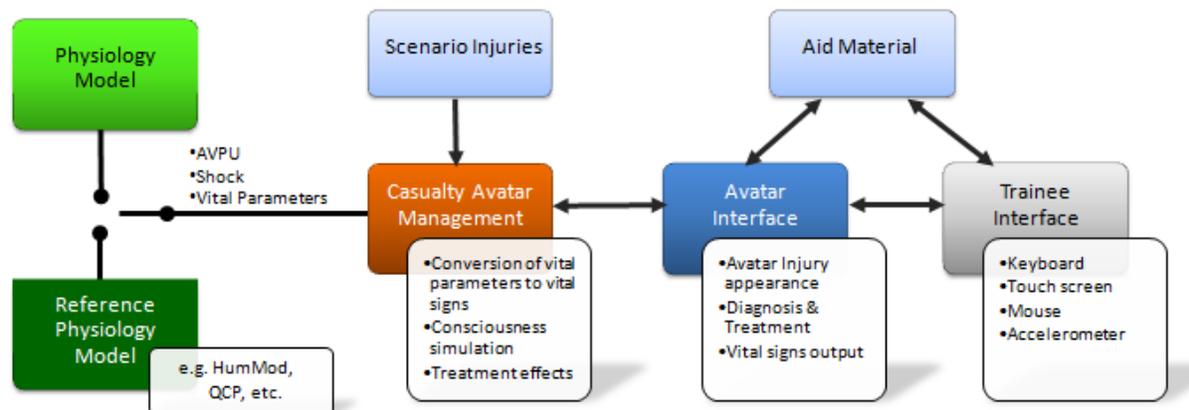


Figure 2: Physiology model comparison and casualty avatar management architecture.

#### 4.3.1 Scenario Injuries, Vital Signs Diagnosis, Treatment and Casualty Response Management

The physiology models only provide time series of vital parameters; ideally also shock and AVPU consciousness information; the additional modules “Casualty Avatar Management” and “Avatar Interface” (see Figure 2) are needed to respectively compute and display casualty reactions and casualty vital signs. These modules need to be individually validated as well.

In order to support multiple casualties and multiple players, the Casualty Avatar Management module detects players within reach of any casualty. This proximity detection causes a casualty diagnosis display and treatment menu to appear on screen close to the casualty representation prepared by the Avatar Interface, offering a variety of menu and submenu choices to the trainee, which may include:

- open the medical supplies backpack (to select an item to use on the casualty),
- move the casualty, including: turning it over to examine its back for pneumothorax diagnosis, transporting it to a safer, tactically less exposed place, or putting it in recovery position,
- diagnosis action selection, including: pulse point selection, blood pressure, heart rate or SPO<sub>2</sub> measurement, communicate [this could open a submenu with a list of questions to evaluate AVPU consciousness], touch or press to elicit a casualty reaction or for diagnosis purposes, remove clothes for pneumothorax examination or bandage application, and

- treatment action including: apply bandage [size and position], press on bandage [pressure], apply tourniquet [location], tighten tourniquet [number of turns], clear airway, apply needle [location], pierce with needle [two angles and one depth], etc.

Upon menu selection by the trainee of a pressure bandage, the Avatar Interface causes it to appear near the casualty, enabling the trainee to graphically position and press it in place as needed. The Avatar Interface then transmits the treatment type and application parameters (position and pressure) back to the Casualty Avatar Management module, which checks them against injury features preset by the scenario (Was a tourniquet needed? Is the location correct?), derives a treatment effectiveness based on those values, computes the changed vital parameters (updated blood volume change) and notifies them to the physiology model.

The Avatar Interface, driven by the Casualty Avatar Management module, manages the sounds and appearance of each casualty (amputation, blood leakage, wound location, facial expressions, speech synthesis, heart and breathing sounds, etc.) and general casualty appearance.

The Casualty Avatar Management module may be implemented as an expert system which interprets the time series of AVPU status, shock status and vital parameters from any of the physiological models and decides/computes on that basis which casualty reactions and vital signs, such as lip color, radial pulse, heart and breathing noises will be transmitted or not (no radial pulse in case of shock) to the Avatar Interface to be shown to the trainee. These reactions may include screams, movements, eyelid or pupil movement, consciousness manifestations including speech contents, quality and tone, expressing pain location or intensity, anger, anxiety, confusion, grunts, labored breathing. The Casualty Avatar Management module further initialises or reinitialises as necessary both the physiology model and the casualty representation according to the specific injuries (type, location, severity) foreseen by the scenario, resp. as updated by stimuli or treatment types and locations applied by the trainee to the casualty.

#### **4.3.2 Face Validation of the Casualty Avatar Management Module**

The complexity of the Casualty Avatar Management module heavily depends on the expectations of the training authorities, of the game experts and of the platform capabilities. For instance, the consciousness simulation can be carried out quite simply as answer to the “communicate” command in form of a text message: “from the casualty's response you determine that his AVPU consciousness status is P” or in a more realistic and complex manner by trying to respond with actual speech in a way which permits a clear diagnosis. Clearly the second option is superior, but much more difficult to validate and troubleshoot.

Another validation challenge has been the tourniquet implementation, essential for TCCC training: Tourniquets are used to stop massive bleeding on extremities. For these types of injuries they are in general the best aid action during the Care under Fire phase on the battlefield. In a first implementation, placing the tourniquet icon on the casualty near the injury instantly stopped bleeding, based on the principle that the intended purpose of the game is to teach cognitive diagnosis and treatment decision processes, not manual skills like bandaging or tourniquet application which are best learnt on manikins or the trainee's own limbs. The game thus validly taught that tourniquets are suitable for such injuries. However, a German army TCCC expert objected that tourniquets frequently are not sufficiently tightened by trainees, and that bleeding may restart even in cases the tourniquet was correctly installed. They insisted on a less simplistic model with more realistic tightness adjustment, for the purpose of teaching trainees that injuries must constantly be rechecked and tourniquets may have to be readjusted from time to time. Since the effectiveness of a tourniquet depends on the blood pressure and movements of the casualty, our new approach is based on a blood pressure model. After several rounds of internal and external face validation the model is now able to simulate the effects of various causes on blood pressure and bleeding, and consequently, the effects of blood pressure on the effectiveness of the tourniquet. The most recent step included the possibility to improvising a tourniquet with a belt.

## 5 CONCLUSION

Facing the task of designing a TCCC serious game demonstrator and mindful of the substantial medical risks and ethical challenges posed by this particular category of training methods intended to teach life-saving behavior, an identification and evaluation of best practice in this field has been attempted. As the evaluation of a TCCC serious game as a training method according to APA guidelines must be postponed until completion of the first prototype of the game, a preliminary validation of the simulation components during the early phases of development, particularly a validation of its pathophysiological model, is essential, but is critically limited by the near-impossibility of obtaining reliable experimental data in parts of the domain of applicability due to ethical limitations on life-threatening medical experimentation.

To solve this validation dilemma, the modeled vital parameter trajectories of virtual casualties should be compared with the outputs of other approved physiology models, textbook or expert knowledge and theoretical findings. Appropriate approaches have been proposed to meet these requirements, reflecting the current state of the art, describing possible reference models, in particular strongly calibrated integrative physiology models used for teaching purposes in medical schools to obtain surrogate validation data, and defining a framework for a more rigorous validation process not relying solely on subjective criteria.

Further validation challenges requiring further study include the difficulty to capture the variability of human physiology and to define exactly the limits of applicability of TCCC simulations, the pedagogical pros and cons of reproducing this variability, the importance of plausibility as a weak surrogate of exact validity etc.

We are still at the beginning of our demonstrator design effort and hope to gain further insight in its validation as we go along: there are still a few years of work before actual comparative training effectiveness validation results can be presented.

## ACKNOWLEDGMENTS

The authors wish to thank the Sanitätsamt der Bundeswehr (esp. Dr. L. Schneidereit) for funding this research, as well as the project contributors and partners, such as Maj. Florent Josse (medical expert from Bundeswehr), and many colleagues at Institut für Technik Intelligenter Systeme (ITIS GmbH), at the Universität der Bundeswehr and elsewhere, for their help, encouragements and valuable contributions.

## REFERENCES

- Abram, S. R., B. L. Hodnett, R. L. Summers, T. G. Coleman, and R. L. Hester. 2007. "Quantitative Circulatory Physiology: An Integrative Mathematical Model of Human Physiology for Medical Education." *Advances in Physiology Education* 31:202-210.
- Anton, J. J., E. Burns, J. Norfleet, and D. Thompson. "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 *Proceedings of the 2009 ITEA Live-Virtual-Constructive Conference: Requirements for Modeling and Simulation in Test & Evaluation*. El Paso, TX.
- Butler, F. K. 2007. "Tactical Combat Casualty Care 2007: Evolving Concepts and Battlefield Experience." *Military Medicine* 172(11 Suppl):1-19.
- Coleman, T. G., and J. E. Randall. 1983. "HUMAN. A Comprehensive Physiological Model." *Physiologist* 26:15-21.
- Dancy, C. L., F. E. Ritter, and K. Berry. "Towards Adding a Physiological Substrate to ACT-R." In *Proceedings of the 2012 21st Conference on Behavior Representation in Modeling and Simulation*, 78-85. Amelia Island, FL.
- Feron, H., and M. Hofmann. "Tactical Combat Casualty Care: Strategic Issues of a Serious Simulation Game Development." In *Proceedings of the 2012 Winter Simulation Conference*, edited by C.

- Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers (IEEE), Inc.
- Gallagher, A. G., E. M. Ritter, and R. M. Satava. 2003. "Fundamental Principles of Validation and Reliability: Rigorous Science for the Assessment of Surgical Education and Training." *Surgical Endoscopy and other Interventional Techniques* 17(10):1525-1529.
- Graafland, M., J. M. Schraagen, and M. P. Schijven. 2012. "Systematic Review of Serious Games for Medical Education and Surgical Skills Training." *British Journal of Surgery* 99:1322-1330.
- Guyton, A. C., T. G. Coleman, and H. J. Granger. 1972. "Circulation: Overall Regulation." *Annual Review of Physiology* 34:13-46.
- Hester, R. L., R. Iliescu, R. Summers, and T. G. Coleman. 2011a. "Systems Biology and Integrative Physiological Modeling." *The Journal of Physiology* 589(5):1053-1060.
- Hester, R. L., A. Brown, L. Husband, R. Iliescu, W. A. Pruet, R. L. Summers, and T. Coleman. 2011b. "HumMod: A Modeling Environment for the Simulation of Integrative Human Physiology." *Frontiers in Physiology* 2:12.
- Hester, R. L., P. Kenny, T. Parsons, A. Rizzo, T. Amundsen, J. Fu, and T. Coleman. "Implementation of a High-Fidelity Physiological Model in a Virtual Patient." In *Proceedings of the 2012 International Meeting on Simulation in Healthcare (IMSH)*. Wheaton, IL: Society for Simulation in Healthcare.
- Hodges, J. 2012. "Human Factors & Medical Panel Meeting on Advanced Training Technologies for Medical Healthcare Research Technology Group 215 (NATO HFM-RTG-215)." Workshop Proceedings. London, UK.
- Hoitz, J. 2012. "Simulation in Military Medicine: More Than Just Skill Training." *Medical Corps International Military Forum*. Accessed March 20, 2013. [http://www.mci-forum.com/category/experiences/419-Simulation\\_in\\_military\\_medicine\\_more\\_than\\_just\\_skill\\_training.html](http://www.mci-forum.com/category/experiences/419-Simulation_in_military_medicine_more_than_just_skill_training.html).
- Kofranek, J., S. Matousek, J. Rusz, P. Stodulka, P. Privitzer, M. Matejak, and M. Tribula. 2011. "The Atlas of Physiology and Pathophysiology: Web-based Multimedia Enabled Interactive Simulations." *Computer Methods and Programs in Biomedicine* 104(2):143-153.
- MEDIC. 2009. "Medical Education and Demonstration of Individual Competence (MEDIC)." US Department of Army Training Circular (TC) No. 8-800.
- PHTLS. 2010. "PHTLS: Prehospital Trauma Life Support." Military 7<sup>th</sup> Edition. NAEMT. St. Louis, MO: Mosby JEMS/Elsevier.
- Rodríguez-Barbero, A., and J. M. López-Novoa. 2008. "Teaching Integrative Physiology Using the Quantitative Circulatory Physiology Model and Case Discussion Method: Evaluation of the Learning Experience." *Advances in Physiology Education* 32:304-311.
- RTO-TR-HFM-128. 2009. "Human Behavior Representation in Constructive Simulation." RTO Technical Report No. RTO-TR-HFM-128 AC/323(HFM-128)TP/250, Research and Technology Organisation (RTO), North Atlantic Treaty Organisation (NATO), Neuilly-sur-Seine, France. <http://ftp.rta.nato.int/public/PubFullText/RTO/TR/RTO-TR-HFM-128//SSTR-HFM-128-ALL.pdf>. [Accessed March 20, 2013].
- Schlesinger, S., R. E. Crosbie, R. E. Gagné, G. S. Innis, C. S. Lalwani, J. Loch, R. J. Sylvester, R. D. Wright, N. Kheir, and D. Bartos. 1979. "Terminology for Model Credibility." *Simulation* 32(3):103-104.
- Sotomayor, T. M., and D. Proctor. 2009. "Assessing Combat Medic Knowledge and Transfer Effects Resulting from Alternative Training Treatments." *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 6(3):121-134.
- Sotomayor, T. M., A. M. Salva, and B. W. York. "Measuring the Training Effectiveness of Combat Life-saver Simulation Training Systems." In *Proceedings of the 2012 Interservice/Industry Training, Simulation, and Education Conference (IITSEC)*. Arlington, VA: National Training Systems Association.

- Sotomayor, T. M., B. Quintero, A. M. Salva, and J. Sherrill. "High Fidelity Physiological Model for Immersive Simulation and Training." In *Proceedings of the 2012 Interservice/Industry Training, Simulation, and Education Conference (IITSEC)*. Arlington, VA: National Training Systems Association.
- Summers, R. L., K. Ward, T. Witten, V. Convertino, K. Ryan, T. G. Coleman, and R. L. Hester. 2009. "Validation of a Computational Platform for the Analysis of the Physiologic Mechanisms of a Human Experimental Model of Hemorrhage." *Resuscitation* 80(12):1405-1410.
- TC3Sim. 2010. "Tactical Combat Casualty Care Simulation." Engineering & Computer Simulations, Inc., Orlando, Florida. Accessed March 20, 2013. <http://www.ecsosl.com/solutions/tactical-combat-casualty-care-simulation>.
- TCCC. 2012. "TCCC Guidelines." Military Health System, U.S. Department of Defense. [http://www.health.mil/Libraries/120917\\_TCCC\\_Course\\_Materials/TCCC-Guidelines-120917.pdf](http://www.health.mil/Libraries/120917_TCCC_Course_Materials/TCCC-Guidelines-120917.pdf). [Accessed March 20, 2013].
- TraumaCon. 2012. "Trauma Connect/Combat Lifesaver Mobile Medic Application(MMA)." Engineering & Computer Simulations, Inc., Orlando, Florida. Accessed March 20, 2013. <http://www.ecsosl.com/solutions/traumacon-combat-lifesaver-mma>.
- TREMA. 2012. "Richtlinien der TREMA e.V. für TCCC (Taktischen Verwundetenversorgung)." V.1.2. Tactical Rescue & Emergency Medicine Association e.V., Ulm, Germany. <http://www.tremaonline.info/TREMA%20e.V.%20Guidelines%20fuer%20TCCC%201.2.pdf>. [Accessed March 20, 2013].
- VBS2. 2012. "Virtual Battlespace 2." Bohemia Interactive Australia Pty Ltd. Accessed March 20, 2013. [http://manuals.bisimulations.com/vbs2/2-00/devref/#Adding\\_Models/AM\\_Adding\\_Content.htm](http://manuals.bisimulations.com/vbs2/2-00/devref/#Adding_Models/AM_Adding_Content.htm).
- Verdaasdonk, E. G., L. P. Stassen, L. J. Monteny, and J. Dankelman. 2006. "Validation of a New and Simple Virtual Reality Simulator for Training of Basic Endoscopic Skills: the SIMENDO." *Surgical Endoscopy and Other Interventional Techniques* 20(3):511-518.
- vMedic. 2012. "Virtual Medic Simulation." Engineering & Computer Simulations, Inc., Orlando, Florida. Accessed March 20, 2013. <http://www.ecsosl.com/solutions/vmedic>.
- VMSTC. 2010. "Virtual Medical Simulation Training Center." Engineering & Computer Simulations, Inc., Orlando, Florida. Accessed March 20, 2013. <http://www.ecsosl.com/solutions/virtual-medical-simulation-training-center-vmstc>.

## AUTHOR BIOGRAPHIES

**HWA FERON** is a 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, she obtained a M.S. in artificial intelligence at Illinois Institute of Technology, Chicago. Her e-mail address is [hwa.feron@unibw.de](mailto:hwa.feron@unibw.de).

**MARKO HOFMANN** has studied computer science in Munich. He is a project manager at Institut für Technik Intelligenter Systeme and a lecturer at the University of the Federal Armed Forces Munich, Germany. He has written his dissertation (2000) and "habilitation" (2010) on military computer simulation. His email address is [marko.hofmann@unibw.de](mailto:marko.hofmann@unibw.de).

**AXEL LEHMANN**, Chair Professor of Modelling and Simulation at the University of the Federal Armed Forces Munich, Germany until 2011, is Executive Director of Institut für Technik Intelligenter Systeme. He is a Fellow of the German Informatics Society (Gesellschaft für Informatik e.V.). His email address is [axel.lehmann@unibw.de](mailto:axel.lehmann@unibw.de).