

Designing a Digital Medical Management Training Simulator Using Distributed Cognition Theory

Jonas Rybing, Erik Prytz, Johan Hornwall, Helene Nilsson, Carl-Oscar Jonson and Magnus Bang

Journal Article



N.B.: When citing this work, cite the original article.

Original Publication:

Jonas Rybing, Erik Prytz, Johan Hornwall, Helene Nilsson, Carl-Oscar Jonson and Magnus Bang, Designing a Digital Medical Management Training Simulator Using Distributed Cognition Theory, *Journal Simulation & Gaming*, 2016.

<http://dx.doi.org/10.1177/1046878116676511>

Copyright: SAGE Publications (UK and US)

<http://www.uk.sagepub.com/home.nav>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-132723>

Designing a digital medical management training simulator using distributed cognition theory

Structured abstract

Background: Training of medical professionals is important to improve care during mass-causality events. Therefore, it is essential to extend knowledge on how to design valid and usable simulation-based training environments.

Purpose: This article investigates how distributed cognition and simulation theory concepts can guide design of simulation-based training environments. We present the design and user evaluation of DIGEMERGO, a simulator for training and assessing emergency medicine management.

Design approach: A prior Distributed Cognition in Teamwork (DiCoT) analysis of the Emergo Train System (ETS) guided the design process. The design objective of DIGEMERGO was to be useful, usable, retain distributed cognition features of ETS, and strengthen validity and output reliability.

Evaluation: Eight expert ETS instructors participated in a formative system evaluation. The Technology Assessment Model (TAM) questionnaire was used to measure usefulness and ease of use. Observations and post-test interviews were conducted to contextualize the measures.

Results: The results showed that DIGEMERGO was perceived as somewhat to quite useful and somewhat easy to use. Overall, expert users considered DIGEMERGO promising and successful in retaining core ETS features.

Conclusions: The study indicates that a design methodology based on distributed cognition and simulation theory can be successfully combined to guide simulator (re)design and strengthen simulator validity.

Keywords: Simulation; Simulator design; Distributed cognition; Simulator user evaluation; Emergency medicine management training

Introduction

In the medical domain, major incidents are rare but critical events that must be managed effectively and efficiently. A major incident is defined as a situation in which available everyday response resources are inadequate in relation to the urgent needs (Cohen et al., 2013; Nilsson, 2013). In such situations effective management is essential to minimize negative patient outcomes (Blackwell & Kaufman, 2002). Problems and shortcomings in the management response to disasters and major incidents have been identified, for example, in terms of establishing leadership and roles, managing resources, scaling up the response, and distributing casualties (see e.g. Juffermans & Bierens, 2010; Nilsson, 2013). Many of these problems are due to inappropriate training and education for situations where routine knowledge and everyday experience is insufficient (Bolling et al., 2007; Juffermans & Bierens, 2010; Nilsson, 2013).

Simulation-based training in the form of full-scale exercises with a high degree of realism is one way to address these problems. Such full-scale exercises combined with a structured evaluation process that includes both process- and outcome indicators are often considered the best approach to training (Cohen et al., 2013; Gryth et al., 2010). However, the effects of full-scale medical training are difficult to measure, and the exercises themselves are expensive, complex, and time consuming to arrange. For these reasons, the major part of medical management training is today performed using other techniques, such as low-fidelity simulations (Bewley & O'Neil, 2013; Legemaate, Burkle Jr., & Bierens, 2012; Lennquist, 2005; Nilsson, 2013; Sundnes & Birnbaum, 2003). These types of simulations have practical benefits, such as saving cost and time as well as offering a safe learning environment. Such simulation-based training has been shown to improve learner commitment, support repetitive practice, target specific skills, and support adjustment of training difficulty to match learner competence (Alessi, 1988; Bewley & O'Neil, 2013; Issenberg, McGaghie, Hart et al., 1999; Issenberg, McGaghie, Petrusa, Gordon, & Scalese, 2005). However, to train the correct skills efficiently it is important to simulate the correct things in a useful and repeatable manner adapted to the target users. Currently, a large body of the published research on simulation-based training lacks the quality and rigor to provide a useful foundation for meta-reviews (McGaghie, Issenberg, Petrusa, & Scalese, 2006). Therefore, it is imperative to develop training simulators using a systematic and scientifically rigorous method, so that they can be used in future studies on simulation-based training effectiveness.

Emergo Train System (ETS) is an internationally-used low-fidelity pedagogical simulation tool for training and evaluating emergency medicine management, organizational capacity, and regional medical surge capacity (Nilsson, Vikström, & Rüter, 2010; Nilsson, 2013; Rådestad et al., 2012). ETS employs an analogue user interface (whiteboards, magnetic figures, adhesive labels, whiteboard pens, etc.) to simulate a medical response scenario including the flow of patients from accident site to receiving hospitals (see Figure 1). Patients and resources are

represented by figures placed on whiteboards that represent locations or processes (e.g., on-going transports) in a disaster scenario. As in real life, the participants' task in an ETS simulation is to save lives by making critical decisions and providing necessary treatments according to the patients' needs.



Figure 1: The analogue Emergo Train System (ETS) simulator.

In a previous study, several benefits of implementing a digital version of ETS were identified using a distributed cognition analysis (Rybing et al., 2016). That study also indicated that, although ETS provided flexibility and customizability of scenarios before and during training sessions, the flexibility of the analogue approach made the simulation vulnerable to validity and outcome reliability threats.

The goal of the present study was to design a useful and usable digital prototype version of ETS that would retain essential features and pedagogical aspects of the analogue simulation as well as increasing its validity and outcome reliability. Consequently, this article aims to contribute to the understanding of how simulation validity concerns affects design decisions and user interface implementations of training and simulation systems for emergency management. The article is structured as follows. First, we establish central concepts in simulation theory based on the literature. Second, we present DIGEMERGO, a digital version of ETS, together with a report on its design rationale with a discussion on how design features are related to concepts from simulation theory. Third, we present the results from a user evaluation of DIGEMERGO with regards to its usefulness and ease of use. In the remainder of the article, we discuss different dimensions of validity and reliability in relation to simulation design.

Simulation theory

Simulations are frequently used in healthcare and emergency medicine because simulations allow for interactive and immersive activities in which a clinical experience is recreated without risk to patients (Maran & Glavin, 2003). Formally, a *simulation* is the execution of a *model*, which in turn is a *representation* of a physical, mathematical or otherwise logical system, entity, phenomenon, or process (Hancock, Vincenzi, Wise, & Mouloua, 2010; Sokolowski & Banks, 2010). Further, a distinction is made between *simulation* and *simulator*, where the first term refers to the activity and the latter term refers to the medium (the simulation device) that enables the activity through execution of the model.

Andrews and colleagues (1998) categorized simulations in two dimensions, where the first reflects the real/virtual nature of the people and the simulands (the elements being simulated) and the second captures whether the simulation is considered local or distributed. In the first dimension, simulations are classified as *live*, *virtual*, *constructive*, or *hybrid*. Live simulations involve users that use real systems, such as field exercises, whereas in virtual simulations users use non-real systems, for instance flight simulators. In *constructive simulations*, users control both non-real systems and people, such as in war games. *Hybrid simulations* are ways of combining the other three, often using distributed computing environments. The second dimension captures the geographical nature of a simulation: *local* simulations are simulations confined to actual physical locations, while *distributed* simulations involve networks to share information over a variety of locations.

Simulation fidelity

In a training context, simulation *fidelity* can be defined as the *degree of similarity* between the training situation and the real operational situation that is simulated (Maran & Glavin, 2003). Moreover, it is common to divide fidelity into *engineering fidelity*, which is the level of physical similarity of simulands, and *cognitive* or *psychological fidelity*, which refers to the extent that the simulator engages the trainees in representative cognitive activities (e.g. communication, decision-making, planning, etc.) of the real-world tasks (Hochmitz & Yuviler-Gavish, 2011).

It is a common assumption that higher engineering fidelity results in more efficient training. However, studies indicate that a higher level of fidelity does not directly translate into enhanced learning (Alessi, 1988; Feinstein & Cannon, 2002). On the contrary, in many cases it is the departure from realism that contributes to effective knowledge transfer (Alessi, 1988). One reason for this is that simulations can do things that cannot be done in reality. For instance, simulations can replay sequences, tasks can be simplified to highlight key performance aspects or adapted to participants' skill levels and the simulation objectives, and complicating factors can be planned and controlled (Alessi, 1988; Hancock et al., 2010). Researchers have argued that low engineering fidelity and high cognitive fidelity is valuable to have in training simulations focusing on management or command and control skills (Toups, Kerne, Hamilton, & Shahzad, 2011).

Simulation validity and reliability

The validity of a simulator is important for the learning outcome (Issenberg et al., 2005). *Validity* is a complex concept and there are several sub-classifications to consider when determining validity. Feinstein and Cannon (2002) defines four aspects of validity: representational, application (education), internal, and external validity. *Representational validity* reflects how the simulation *represents* the system or process being simulated and whether the representations manage to capture the essence of what is being simulated. This is a form of *functional equivalence* or *believability* of the simulation, but it is not the same as fidelity because representationally-valid simulations do not have to look like reality. *Application validity* has to do with whether the simulation fulfills its purpose. That is, if the simulation achieves its intended effects on those who use it. This often translates into whether measured performance of targeted skills has improved after simulated training exercise, also known as education validity.

Internal validity can be divided into two parts. The first part considers whether the simulation actually reflects the intended real life phenomena and components. That is, what is simulated and in which ways? The second part reflects how participants can understand, interact with, and gain insight from the simulation. Both parts of internal validity are captured by the following question: is the model of reality accurately constructed and correctly executed? *External validity* reflects whether the output and behavior of the simulation is similar and transferable to the modeled system. It differs from application validity in the sense that high application validity does not necessarily imply an accurate behavioral replication of what would happen in real life, but still teaches the skills needed for a real situation. External validity can also be viewed as the degree to which skills and knowledge used in reality are necessary for performing well in the simulation.

Reliability of the simulator and of the simulation outcome is concerned with the replicability of the simulation procedure and the results – outcome – of the simulation. Simulations with a high degree of reliability will produce consistent output from one measurement to another, that is, at different times, using different raters (instructors), and with different (but equivalent) tasks (Bewley & O'Neil, 2013).

Design procedure

An important goal when designing the DIGEMERGO simulator was to retain the pedagogical method and user interaction model of the analogue ETS environment and increase simulation validity and reliability. To achieve this goal, we employed a two-stage design method. First, we analyzed ETS in depth using a cognitive task analysis, as recommended by van Engen-Verheul and colleagues (2016). This analysis also generated a list of potential benefits and drawbacks

of implementing a digital version. Second, we summarized and categorized identified cognitive features of ETS with the goal to find a set of *core qualities* from a distributed cognitive perspective that characterize ETS simulations. The intention was to retain these important core qualities - the ETS *essence* - in DIGEMERGO. The features are shown in Table 1. This analysis is reported in full elsewhere (Rybing et al., 2016), but the key points are summarized below.

The cognitive task analysis employed the Distributed Cognition in Teamwork (DiCoT; Furniss & Blandford, 2006) methodology to analyze ETS from a *distributed cognition* perspective (Hutchins & Lintern, 1995; Perry, 2003). Distributed cognition theory (DCog) views cognition as something that is shaped through a cognitive agent's (e.g., a person's) interaction with the environment, with tools and artefacts, and with other agents in a *cognitive system*. Cognitive processes are in this sense distributed over the entire cognitive system (Hutchins & Lintern 1995). DCog theory has been used to study how tools and artefacts shape, for instance, decision making or how they are used to offload cognitively demanding tasks (e.g. through their spatial arrangement, see for instance Kirsh, 1995, or Perry, 2003). DCog theory can therefore be a valuable technique for studying how the structure of a simulator shapes the cognitive processes involved in the task at hand, and how the simulator design will affect aspects of validity and fidelity that involve cognitive tasks, such as emergency medicine management. DiCoT, on the other hand, is a methodological framework grounded in DCog theory for studying teamwork with a focus on how tools and resources in the environment (in our case the ETS simulator) supports cognitive teamwork processes. Typically, a DiCoT study comprises ethnographic fieldwork where the researcher or designer observes cognitive processes through artefact-use, which is analyzed through the lens of distributed cognition theory according to a set of principles. A DiCoT-analysis is also intended to be useful for (re)design of cognitive artefacts and systems, as it can highlight areas where the design of these could be improved (Furniss & Blandford, 2006). The result of our DiCoT analysis in the previous study (Rybing et al., 2016) was a description of cognitive functions, tasks, artefacts, and work processes of ETS expressed in the vocabulary of distributed cognition.

Table 1: Core features of ETS identified with DiCoT.

ETS Feature	Example of feature in action
<i>Representational openness.</i> Openness ¹ (viewability) of simulands allow learners use whiteboards as information hubs and as means of sharing information, contributing to a shared situation awareness. Instructors also profit from openness by gaining an overview of the exercise state, helping	Learners often group (spatially co-locate) patient figures on the whiteboards to communicate which hospital the patients are to be transported to.

¹ See Garbis (2000) for more information on artefact openness.

<p>them to maintain the situation awareness required to manage the exercise.</p>	
<p><i>Representational and interaction abstraction.</i> Patient medical status is represented by a few lines of text on a magnetic figure card. These low-fidelity representations and available interactions with magnetic patient figures (see Figure 1) allow for medical decision making, not for medical practice. Learners interact with and use the representations by moving and grouping them on the whiteboards. In doing so they offload cognition, transform tasks, and share information.</p>	<p>Typically, when treating or triaging a patient the user typically looks at, picks up, holds, flips, and reads the medical information on the front and back of the figure card. When a medical decision is made, the treatment or triage is performed by attaching an adhesive label on the patient figure.</p>
<p><i>Representational customizability.</i> Learners and instructors frequently annotate and highlight specific figures, or figure features, by marking the figures themselves or the surrounding whiteboard space. This allows for offloading of cognition, transformation of tasks, and sharing of information.</p>	<p>Learners often write next to, or sometimes directly on, the figure the time at which an ongoing treatment will finish.</p>
<p><i>Simulation scalability, variety, and flexibility.</i> The representations used enable scenarios to vary in size, ranging from a few participants to over twenty, and nature. Different types of scenarios can be simulated, including both prehospital and in-hospital emergency management. Further, scenarios are modifiable during play – leaving room for both participant and instructor improvisations. The flexible use of the simulation artefacts can trigger different types of learner behavior.</p>	<p>Learners sometimes ask for resources that are not represented in ETS by a specific magnetic symbol. In such cases, instructors can verbally acknowledge the resources existence, and learners can behave in accordance to this without “breaking the rules”.</p>
<p><i>Use of instructors.</i> The instructors play a big part in controlling the simulation, measuring learner performance, enforcing simulation rules, assisting and teaching learners, and contributing to the immersion of the simulation by acting as confederates.</p>	<p>Managing the simulation involves executing a written scenario script, keeping track of patients and their transport and treatment times, calculate outcome and perform after action review.</p>

<p><i>Simulation output measures.</i> The most important output measure of learner performance is that of patient outcome, i.e. whether a patient is at risk of preventable complications or death. Essentially, three main components in the ETS model are measured (by instructors) during a simulation to determine patient outcome: patient triage, patient treatment, and time until hospital care was received. This feature is highly dependent on and determined by other features, such as the representational openness and representational customizability.</p>	<p>Instructors measure time by using a smartphone app or a wall clock (also observable by the participants) and record it on a whiteboard or a laptop. Timing all transport periods and scenario events was seen as the most demanding task for an instructor.</p>
---	--

Further, the DiCoT analysis enabled us to critically analyze how phenomena occurring in the ETS simulation could lead to issues in terms of validity and output reliability. For instance, overhearing information (unintended information movement) could lead to unrealistic situation awareness amongst the learners. Another issue identified was inconsistencies in adhering to treatment and transport times as well as resource requirements. Such validity issues could skew performance output measures, making them less reliable. Table 2 provides a summary of such validity issues.

Table 2: Identified validity and reliability issues in ETS.

Identified Issue	Issue type
<p><i>Unrealistic situation awareness.</i> Learners might oversee and overhear information that should not be available to them. For example, learners were observed to anticipate events, as they could see instructors preparing the materials. Relates to feature: <i>representational openness</i>.</p>	<p>Internal, external</p>
<p><i>Corner cutting.</i> The interaction style and abstracted actions of ETS makes it easy for learners to sidestep rules of the simulation. For example, when treating a patient, learners frequently ‘cut corners’ in terms of not performing required actions or not adhering to time required for the action. This causes unrealistic treatment times and performance scores. Relates to features: <i>representational and interaction abstraction; simulation scalability, variety, and flexibility</i>.</p>	<p>Internal, external, reliability</p>

<p><i>Static simuland states.</i> All simulands' (patients and resources) states are static and unchangeable, which for example means that the health state of a patient does not change over time or in response to treatment. Relates to feature: <i>representational and interaction abstraction.</i></p>	<p>Internal, external, representational</p>
<p><i>Subjectiveness.</i> The flexibility of the simulator blurs which actions are allowed, leading to subjective decision by the instructor or learner. This can however be used as feature in training. Relates to features: <i>representational and interaction abstraction, simulation scalability, variety, and flexibility; use of instructors</i></p>	<p>Internal, external</p>
<p><i>Manual control.</i> Manual simulation control (timing, enforcing rules, moving figures, etc.) is prone to estimations, errors, and slips - causing simulation unreliability. Relates to features: <i>representational and interaction abstraction; use of instructors</i></p>	<p>Internal, reliability</p>
<p><i>High instructor load.</i> Manually controlling the simulation and gathering data for after-action-review occupies an instructor's time and cognitive resources that otherwise could be spent on assisting learners or acting as confederates. Relates to features: <i>representational and interaction abstraction; use of instructors; simulation output measures.</i></p>	<p>Application</p>
<p><i>Patient outcome measures.</i> The patient outcome is calculated based on the time at which a patient reached a care destination or received a treatment. These measures are vulnerable to the issues <i>cutting corners</i> and <i>instructor load</i>. Further, the outcome measures are simplified to enable instructors to quickly apply them. Relates to features: <i>representational and interaction abstraction; use of instructors; simulation output measures.</i></p>	<p>Reliability, internal</p>
<p><i>Setup and replay time.</i> Setting up and resetting scenarios is ETS takes time and recourses, making it difficult to conduct consecutive simulations. Relates to features: <i>representational and interaction abstraction; use of instructors.</i></p>	<p>Application</p>

The abstracted interaction and representation style of ETS, as reported in Table 1, indicates that this simulator has low engineering fidelity. However, this property together with representation openness and customizability enables a “walk-up-and-use interface” that place focus on working *at* the interface rather than working *with* the simulation interface. What this means is that a user does not have to be skilled at using the simulator to perform well in the simulation. The walk-up-and-use interface moves emphasis from technical medical skills (performing treatments) of the simulation to strategic management skills (medical decision

making, communication, etc.), which is appropriate as the simulator is used for training and evaluating medical management - not medical practice. Therefore, cognitive validity is important for the simulator. However, a 'scalable' engineering fidelity could increase and diversify the applicability of the simulator. On the negative side, the walk-up-and-use interface and the overall analogue structure of ETS is hard to control and demanding for instructors, and this can lead to loss of simulation validity and simulation reliability in terms of simulation output measures. Moreover, as can be seen in Table 2, most of these issues are related to the abstracted interaction and representation style and the roles of the instructors necessary to run the simulation.

Together, the features of ETS in Table 1 is and the identified issues associated with ETS in Table 2 was used as design guidelines during the iterative prototype (re)design and development process of the digital prototype.

Digital simulator

The DIGEMERGO system consists of a server (the main program), clients (user interfaces), simulation managers (main instructor control tool), and an editor used for creating scenarios (see Figure 2). All these components run on standard Windows network-enabled computers. Clients connect remotely to the server and users interact with the system via large multi-touch wall screens (i.e., "digital whiteboards"). Since the system is network-based, it is also possible to connect to the server to observe and monitor the progress of a training session remotely. The knowledge and rules of the simulation, regarding, for instance how treatments are allowed to be administered, are stored in a central database on the server.

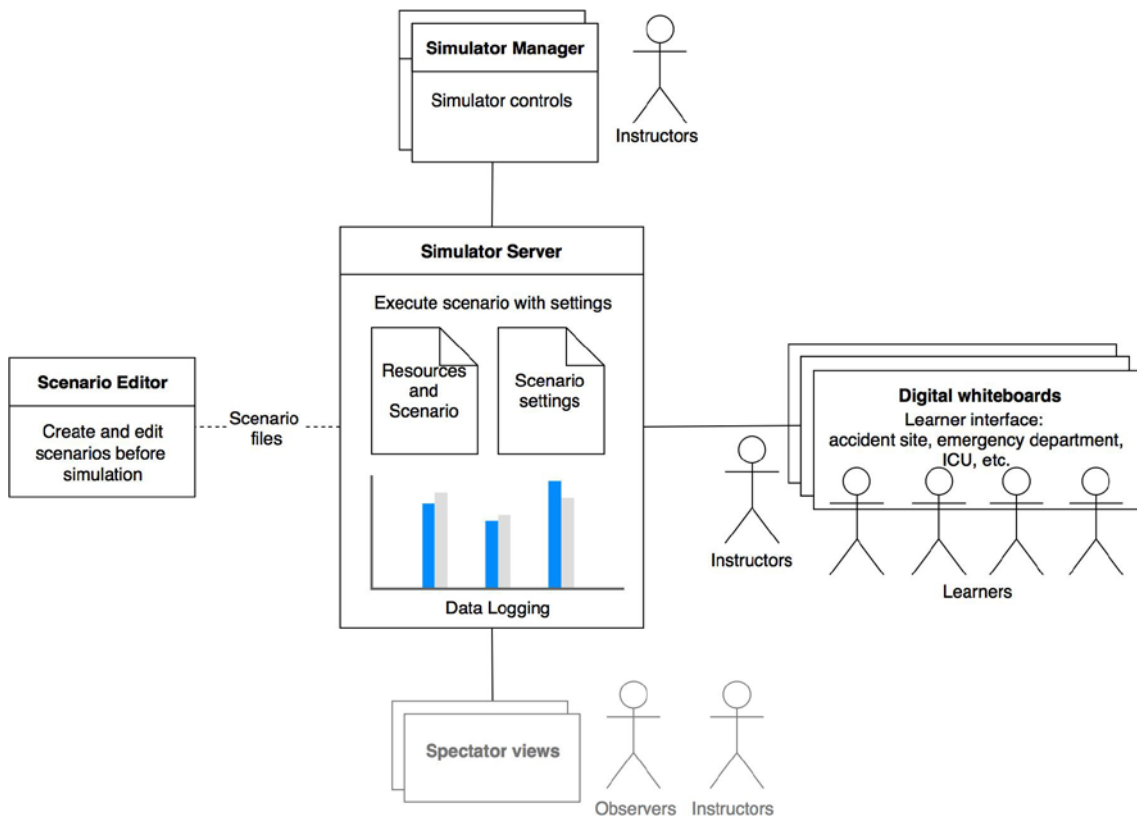


Figure 2: DIGEMERGO simulator architecture overview.

The proposed design solution aimed to retain many ETS features by employing large multi-touch screens, enabling users to interact with digital figures that share the same look as their analogue counterparts (see Figure 3). In this way, DIGEMERGO adopts many functional and visual metaphors from ETS - including *skeumorphisms* - to preserve familiarity. A skeumorph is an attribute that has little intrinsic functionality beyond recalling recognizable features from an earlier system, such as having a trash can icon on the digital desktop (McAulay, 2007). According to Cass and Lauer (2004), skeumorphs act as conceptual bridges between two different technologies. Table 3 and 4 describe how the design of DIGEMERGO is intended to retain important features of ETS while at the same time addressing the validity and reliability issues identified.

Table 3: Implemented features of ETS in DIGEMERGO.

ETS Feature	DIGEMERGO implementation
-------------	--------------------------

Representational openness	Large vertical multi-touch screens are used as ‘digital whiteboards’, which provides collaborative overview and openness in the simulation similar to ETS.
Representational and interaction abstraction	Graphical representations used in ETS were adopted as digitalized versions to retain familiarity and the interaction style of ETS, which was found important for ETS application validity. The multi-touch screens allow direct manipulation of the simulated figures from multiple users at the same time. The digital figures can be spatially arranged as in ETS, and grouping is shown explicitly by figures “snapping” together.
Representational customizability	Figures can be moved, spatially aligned, and grouped as in ETS. Digital ink allows for annotation.
Simulation scalability, variety, and flexibility	The prototype allows for different types of simulation scenarios (prehospital, in-hospital) and simulation size (single or multiple screens, single or multiple instructor controls) of scenarios. Instructors can follow pre-specified scenario script of events (e.g. adding/removing simulands such as resources or patients) to lead the simulation, but events can also be improvised on the fly during simulation. Further, rules and strictness of the simulation can be adjusted depending on simulation purpose (e.g. training or evaluation).
Use of instructors	The simulator was designed to maintain the confederate, assisting/teaching, and overall controlling roles of the instructor, but to take over ‘tedious’ parts of the administering role (e.g. data gathering and time logging, time measuring, rule enforcement, simulation setup).
Simulation output measures	The three components to calculate patient outcome (patient triage, patient treatment, and time until hospital care) are automatically logged and timestamped by the prototype, enabling automatic calculation of patient outcome and other more complex simulation output measures. Further, available data enable new visualizations of simulation outcome during after-action-review post simulation.

Table 4: Mitigating issues with DIGEMERGO.

Identified issue	DIGEMERGO implementation
Unrealistic situation awareness	It is possible to geographically separate sensitive simulated locations from each other, as there is no need to manually move figures between them. Thus, the risk of participants at the one location overhearing what goes on at the other can be eliminated.
Corner cutting	Based on the applied rules (levels of strictness), the prototype can prevent certain actions if its prerequisite conditions are not met, such as treating a patient without having the required resources available or transporting a patient before treatment is completed. The instructor can select the level of simulation strictness that a scenario should adhere to.
Static simuland states	Currently patients are static in the digital prototype, but the digital medium allows for implementation of dynamic content in the future.
Subjectiveness	Simulator enforced rules of interaction and automatic measures can avoid parts of the issues related to subjectiveness.
Manual control error	Automatic measuring of simulation data, digital “moving” of simulands, and closely coupled simulation script can reduce risks for faulty estimations, errors, and slips.
Instructor workload	Instructor workload is reduced through automation of data-gathering and rule enforcing tasks, allowing more time for instructors to interact with participants.
Patient outcome	Automatic measuring and built in rule enforcement can increase the validity and reliability of outcome measures. Further, the possibility to automatically gather all measurable data enables new (more complex) types of output measures and outcome visualizations.
Setup and replay	Once the hardware is set up, initiating and resetting a digital scenario for replay takes virtually no time.

By comparing ETS to DIGEMERGO using the terminology of Andrews et al. (1998), we can see that both simulators are similar *in function*. Regarding the first dimension, both simulators are considered *constructive*, for example both the magnet figures and digital figures are “simulated people”. Further, instructors in both simulators are to take confederate roles (e.g.

by acting as part of the simulation), which is also a form of simulated person. In the second dimension, simulations using ETS should be considered a *local* (constrained to actual physical location). There is no such physical constraint in the DIGEMERGO simulator, making its simulations *distributable*.

Formative evaluation

A formative evaluation was performed with a group of users with no experience of DIGEMERGO but with expert knowledge of ETS. The purpose of the evaluation was to investigate and measure *usefulness* and *ease of use* of the prototype, and to guide further development. The ETS expert users were tasked with using the system as ordinary ETS trainees. This setup allowed us to both evaluate the general usefulness of the new system and take advantage of the expert group's rich experience of the core ETS features. Hence, the overall evaluation can be considered as a combination of standard user-based system evaluation and subject-matter expert review. See Jaspers (2009) for a comparison of usability evaluation methods.

Eight certified ETS senior instructors, all of which had a background in nursing, were recruited to participate in the study. One additional non-certified instructor with previous ETS experience was also included in the user group, making the total number of participants nine. All participants had also experience of being ETS trainees.



Figure 3: Evaluation of DIGEMERGO. To the left, one instructor (leftmost person) and three participants assess the initial state of the accident site. To the right, two participants interact with the DIGEMERGO touch interface.

Scenario and evaluation procedure

The evaluation session employed an ETS scenario for pre-hospital medical management at a 'fire in building' accident site (see Figure 3). This scenario is a standard ETS scenario for ambulance personnel training and evaluation, and was therefore considered as an appropriate

benchmark for the digital system. In this scenario, the trainees take the role as ambulance personnel arriving at an accident site with 20 victims in need of medical attention. The participants arrived to the accident site (i.e., entered the room) as the first and second ambulance on site whereas subsequent ambulances that arrived were represented as resource figures in the simulation.

The DIGEMERGO simulator server and client were set up on a single computer connected to a 68-inch SmartBoard multi-touch screen. Hence, using ordinary touch actions, participants could directly manipulate simulands in DIGEMERGO such as patients and transports. Additionally, they could augment and annotate their workspace using digital pens available in the SmartBoard environment. An additional instructor, with experience in using DIGEMERGO, controlled the simulation using the manager software on a laptop PC stationed on the other side of the room - overlooking the SmartBoard. A further trained instructor took on a confederate role, acting as a collaborating rescue services commander.

Three evaluation sessions were conducted, in which participants partook in groups of two to four. Participants were first given a five-minute introduction to the system and subsequently allowed to interact freely with the system and get accustomed to it. When comfortable, participants left the room and were called in again to enact ambulance personnel in the 30-minute 'fire in building' scenario. After the scenario was completed an open discussion session was initiated, giving the participants an opportunity to openly share first impressions.

Data collection

Qualitative and quantitative data was gathered during the evaluation. The data collection consisted of notes from two observers, post-session group interviews, and usability assessment questionnaires. During the simulation sessions, the two observers took notes using pen and paper. After the scenario was completed, an open group discussion was held in the simulator room, giving participants an opportunity to comment on the simulation experience and share their first impressions. Notes gathered by the observers were used to guide these group interviews.

Two types of questionnaires were used in the study; the first form contained a set of qualitative questions (with open-ended free text answers) that aimed to compare the digital simulation to the analogue counterpart; the second questionnaire was a Technology Acceptance Model (TAM) questionnaire (Davis, 1989). TAM is an approach model to anticipate levels of use and acceptance of a system based on hypothesized relationships between *perceived usefulness*, *perceived ease of use* and *actual system use*. The TAM questionnaire consists of 12 statements that are scored using seven-graded Likert scales. The qualitative data from the first questionnaire (free-text answers), post-session group interviews, and observation notes were coded and categorized into the two main categories of perceived usefulness and perceived ease of use. All qualitative segments included in the current analysis

were mentioned by at least two separate users, the quotes are excerpts from the free text answers and the group interviews. All quotes have been translated from Swedish by the authors.

Results

In the following two sections, summaries of the qualitative results are presented together with the quantitative data from the TAM questionnaire.

Perceived usefulness

The participants' ratings in response to perceived usefulness in TAM is presented in Figure 4. The mean across the six questions pertaining to usefulness was 5.1 and the median was 6 (out of 7). The label associated with a score of 5 in TAM is "somewhat useful", and 6 is "quite useful". Most participants in the evaluation perceived the digital system to function as the analogue system and participants were further observed to engage in a similar level of immersion as had been previously observed in ETS (Rybing et al., 2016).

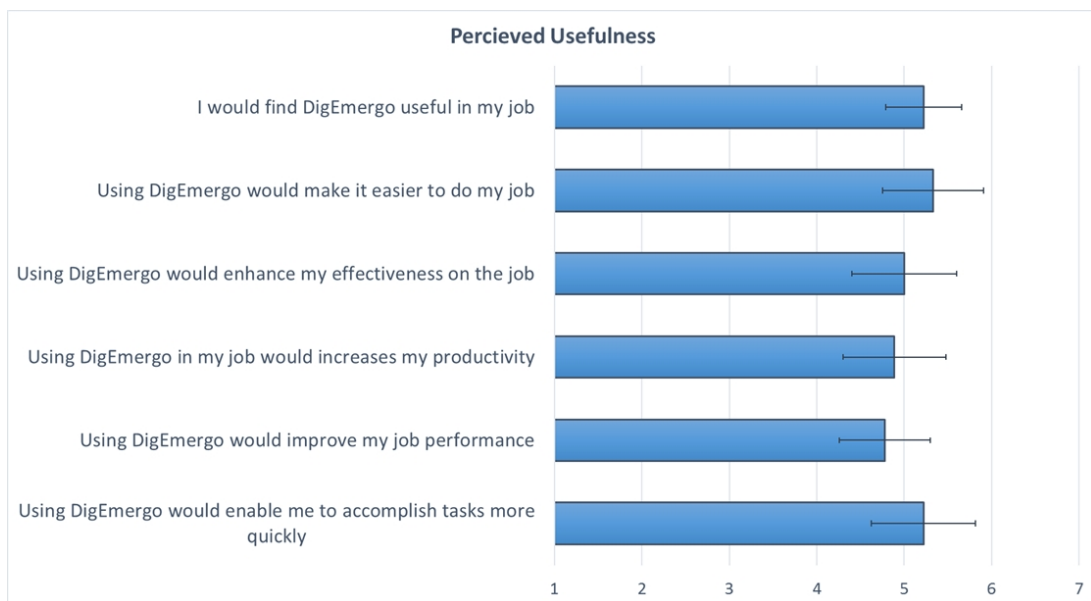


Figure 4. Bar chart of TAM scores including standard error on perceived usefulness of DIGEMERGO.

Several participants stated that using DIGEMERGO would increase control over several aspects of an exercise. In particular, measurement and control of simulation time was frequently

mentioned to be a practical benefit. An example of this is a quote from a participant who highlights the simulator's ability to keep track of treatment times:

"Very good that [treatment] times are stored [in the system]! This simplifies [the simulation] a great deal."

During post-session group interviews, participants mentioned that such tasks occupy much of the instructors' time and cognitive resources in ETS. Furthermore, the participants discussed how automatic measures of treatments and transports times could contribute to a more accurate calculation of patient outcomes. The user group claimed that this feature would lead to increased 'believability' of the simulation. Several participants remarked on the digital system's ability to log data and that this function could enhance after action reviews. Overall, the participants often discussed new applications and features of the digital system in comparison to the analogue counterpart. For instance, participants suggested that the digital system could be useful to perform exercises that are distributed over a large geographical distance – which is hard to coordinate using the analogue system. Even mundane tasks, such as preparation and resetting a scenario, were discussed to be a big benefit in the digital system. This advantage is highlighted by the following quote:

"You don't have to deal with symbols falling off the whiteboard or disappear. Additionally, it becomes much easier both to prepare and to reset an exercise. Major time saver."

The participants also discussed negatives of the system's usefulness. In relation to the flexibility and ease of use of the analogue system, the instructors were concerned that a digital system would reduce the possibility to improvise and adapt during an exercise. Moreover, they were concerned that the digital approach might detach instructors from the participants, losing the close interplay between participants and instructors. The following quote exemplifies both these concerns:

"In ETS, one [the instructor] can improvise based on what the participants are or are not doing. Can this be done in DigEmergo?"

Furthermore, many of the participants mentioned the need to own touch-display hardware to run a simulation – something that is not as common as ordinary whiteboards, a limitation that would make the system less useful.

Perceived ease of use

Figure 5 shows the TAM scores for perceived ease of use. The mean of the six questions relating to ease of use was 4.6 and the median 4.5 (out of 7). The label associated with a score of 5 (rounded) in TAM is "somewhat easy to use". The qualitative data indicated that several participants thought the system was overall both novel and fun to use.

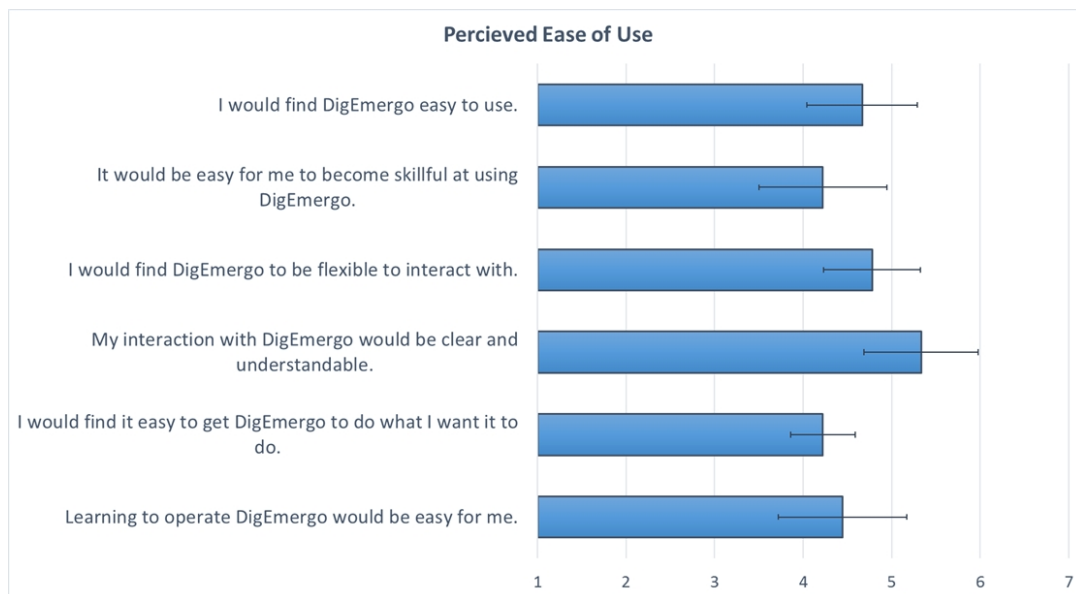


Figure 5. Bar chart of TAM scores including standard error on perceived ease of use of DIGEMERGO.

The participants commented on both benefits and drawbacks with the digital interface. The ability to permanently “flip” patient figures to view their backside (manipulate their displayed information) was seen as something positive that cannot be done in ETS. Furthermore, a perceived drawback was that grouping and manipulating the digital figures was somewhat difficult and that there exists an increased need to train using the system interface – a particular concern for those not familiar with touch-based interaction techniques. One participant expressed this concern as follows:

“The problem could be that the symbols ‘glide into’ each other on the smartboard, which does not happen on the whiteboard.”

Furthermore, a concern was the quality of the hardware and how this will affect system use. For instance, large low-resolution screens are difficult to read when standing close to them. Participants also commented that a scenario, of the size used in the evaluation, in which some 20 patients and 40 resource figures were used, requires multiple screens. It was considered problematic that the single screen used in the evaluation became cluttered. However, despite the lack of workspace on the screen several participants claimed that the digital system provided a good sense of overview, and some hypothesized that this would be especially valuable for the instructors. Furthermore, several participants claimed that the simulator was fun to use. The following quotes are examples of these points:

“The sense of overview makes it easy for instructors and participants to see e.g. that ambulances start a transport and disappears.”

"[DigEmergo is] more fun [than ETS]. Clearer. Treatments are easier to overview."

Discussion

Education and training of medical professionals are needed to improve emergency medicine response. One promising approach is to employ scenario-based simulations where individual performance can be measured in relation to the patient outcomes (Nilsson, 2013). In this article, we presented DIGEMERGO – a prototype simulation system that allows medical professionals to train and evaluate emergency medicine management and medical surge capacity in a digital environment. DIGEMERGO is a re-design and transition from an analogue simulator using a development and design processes that was guided by distributed cognition applied to simulation theory. The study demonstrates how distributed cognition analysis, in this case DiCoT, combined with simulation theory concepts can be used in the design of training simulators - an approach that to the best of the authors' knowledge has not been used elsewhere. The idea behind using distributed cognition for designing simulation environments is that the designer/researcher can analyze cognitive processes in a target real situation to gain an understanding of how these cognitive processes are facilitated in the environment. This can aid and inform designers of which artefacts, representations, and cues are needed to be re-created in simulation to train and assess correctly and effectively. Further, the distributed cognition perspective offers a new way to analyze and inform integral concepts such as validity and cognitive fidelity.

The DIGEMERGO prototype scored a mean score of 5.1 (median 6) out of 7 for usefulness in the user evaluation. This result indicates that the system was perceived as somewhat to quite useful. The qualitative data corroborate this indication, but also stressed the importance of instructor-participant interplay. Furthermore, DIGEMERGO scored 4.6 (median 4.5) out of 7 for ease of use. This is considered acceptable for a prototype, but it also indicates that usability flaws exist in the system. For instance, it was observed and commented by participants that creating figure groups (i.e., attaching figures to one another to indicate resource allocation) was problematic and a source of annoyance – especially when the screen became visually cluttered. Additionally, in relation to ease of use, participants reported that the experience of using DIGEMERGO was similar to that of ETS and that the systems had a similar high level of immersion and engagement. Retaining identified features essential for ETS, such as visual metaphors and skeumorphisms, helped to achieve a feeling of resemblance. However, a common critique of skeumorphisms is that such visual elements can take up unnecessary screen space and that they may be difficult to manipulate – as seen in the case of the visual clutter in DIGEMERGO. We believe that system familiarity aspects must be considered in relation to ease of use and efficient use of screen space, and changing interaction or representation style will affect the validity of the simulation.

Moving to a digital simulation from an analogue counterpart allow for new ways of training and debriefing. For example, the system can facilitate quick mini-training sessions with automatically generated feedback; Lukosch and colleagues (2016) recently demonstrated the value of one such microgame for training. This type of training approach can be a valuable supplement to large and expensive simulation-based team training sessions currently conducted at hospitals worldwide. DIGEMERGO training exercises could also be distributed over large geographical areas, which enable possibilities to arrange larger inter-organizational exercises in which multiple 'locations' and organizations participating in the exercise. With regards to cognitive fidelity versus engineering fidelity, Hochmitz and Yuviler-Gavish (2011) concluded that high cognitive fidelity and high-level engineering training methods have complementary advantages. Thus, combining increased engineering fidelity and cognitive training methods can enhance procedural skills acquisition. A digital system enables new ways to increase engineering fidelity, for instance by using dynamic patient states, which could be beneficial for training and assessing experienced users (Alessi, 1988). Overall, the main advantage of DIGEMERGO is not the technical features themselves but the educational practices they enable, in line with prior research (McGaghie, Issenberg, Petrusa, & Scalese, 2010).

The results of this study indicate that the applied design methodology lead to a prototype that was perceived as useful by expert users, which supports the value of using distributed cognition in design of simulators. However, whether the simulator actually can be said to *function*, which is a requirement for usefulness, must be demonstrated in actual use or by performing a validation study of the system. A validation study could demonstrate DIGEMERGO's ability to capture essential skills for emergency medicine management. One possible approach to achieve this is to compare performance measurements from both emergency medicine management experts and novices performing the same simulation exercises using the digital system. If the experts' performance is significantly better than that of novices', the simulator can be said to exhibit internal and external validity; internal because the simulator can be said to have captured and re-created relevant phenomena of emergency medicine management that experts recognize, and external because the experts' knowledge and skills can be successfully applied.

Conclusion

Within disaster medicine research there has been a call for new training and evaluation methods that can improve emergency management and disaster response. This article describes the design and evaluation of DIGEMERGO, a digital tool for training medical professionals in this domain. The system design aimed to retain aspects of a non-digital legacy training system, particularly the representational validity, while also increasing internal validity and outcome reliability of the simulation. The design approach was based on distributed cognition, simulation theory, and feature extraction of the analogue training environment. The

approach allowed informed design decisions to be made by phrasing simulation validity in terms of distributed cognition. The study found DIGEMERGO to be a useful, immersive, and a promising simulator, which indicates that distributed cognition theory and the DiCoT methodology can be valuable for (re)design of medical training simulators.

Acknowledgements

The authors wish to sincerely thank the ETS instructors and simulation participants for enabling and contributing to this study. The authors would also like to give their thanks to Carl Einarson and Mattias Lantz Cronqvist for their hard and diligent work in implementing the DIGEMERGO system. Finally, the authors wish to thank the anonymous reviewers for providing valuable feedback that improved the quality of this article.

Funding

This work was supported by the the Swedish Civil Contingencies Agency (MSB) [grant number 2011-4957].

References

- Alessi, S. M. (1988). Fidelity in the Design of Instructional Simulations. *Journal of Computer-Based Instruction*, 15(2), 40–47. Retrieved from <http://psycnet.apa.org/psycinfo/1989-13468-001>
- Andrews, D. H., Brown, J., Byrnes, J., Chang, J., & Hartman, R. (1998). *Enabling Technology: Analysis of Categories with Potential to Support the Use of Modeling and Simulation in the United States Air Force*.
- Bewley, W. L., & O'Neil, H. F. (2013). Evaluation of medical simulations. *Military Medicine*, 178(10 Suppl), 64–75. <http://doi.org/10.7205/MILMED-D-13-00255>
- Blackwell, T. H., & Kaufman, J. S. (2002). Response time effectiveness: Comparison of response time and survival in an urban emergency medical services system. *ACADEMIC EMERGENCY MEDICINE*, 9(4), 288–295. <http://doi.org/10.1111/j.1553-2712.2002.tb01321.x>
- Bolling, R., Ehrlin, Y., Forsberg, R., Rüter, A., Soest, V., Vikström, T., ... Brändström, H. (2007). KAMEDO Report 90: terrorist attacks in Madrid, Spain, 2004. *Prehospital and Disaster Medicine*, 22(3), 252–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17894223>
- Cass, K., & Lauer, T. W. (2004). Media transitions: The cases of digital imagery and e-mail. *Information Technology & People*, 17(3), 252–267. <http://doi.org/10.1108/09593840410554175>
- Cohen, D., Sevdalis, N., Taylor, D., Kerr, K., Heys, M., Willett, K., ... Darzi, A. (2013). Emergency preparedness in the 21st century: Training and preparation modules in virtual environments. *Resuscitation*, 84(1), 78–84. <http://doi.org/10.1016/j.resuscitation.2012.05.014>
- Davis, F. (1989). Perceived Usefulness, Perceived ease of use, and user acceptance. *MIS Quarterly*, 13(3), 319–339. <http://doi.org/10.2307/249008>
- Feinstein, A. H., & Cannon, H. M. (2002). Constructs of simulation evaluation. *Simulation & Gaming*, 33(4), 425–440. <http://doi.org/10.1177/1046878102238606>
- Furniss, D., & Blandford, A. (2006). Understanding emergency medical dispatch in terms of distributed cognition: A case study. *Ergonomics*, 49(12-13), 1174–203. <http://doi.org/10.1080/00140130600612663>
- Gryth, D., Rådestad, M., Nilsson, H., Nerf, O., Svensson, L., Castrén, M., & Rüter, A. (2010). Evaluation of medical command and control using performance indicators in a full-scale, major aircraft accident exercise. *Prehosp Disaster Med*, 25(2), 118 – 124. Retrieved from <http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=8559047>

- Hancock, P., Vincenzi, D., Wise, J., & Mouloua, M. (2010). *Human factors in simulation and training*. Boca Raton, FL, USA: CRC Press, Inc.
- Hochmitz, I., & Yuviler-Gavish, N. (2011). Physical fidelity versus cognitive fidelity training in procedural skills acquisition. *Human Factors*, 53(5), 489–501.
<http://doi.org/10.1177/0018720811412777>
- Hutchins, E., & Lintern, G. (1995). *Cognition in the Wild*. MIT press: Cambridge, MA.
- Issenberg S, McGaghie WC, Hart IR, et al. (1999). Simulation technology for health care professional skills training and assessment. *JAMA*, 282(9), 861-866.
- Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., Lee Gordon, D., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*, 27(1), 10–28.
<http://doi.org/10.1080/01421590500046924>
- Jaspers, M. W. M. (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. *International Journal of Medical Informatics*, 78(5), 340–353. <http://doi.org/10.1016/j.ijmedinf.2008.10.002>
- Juffermans, J., & Bierens, J. (2010). Recurrent medical response problems during five recent disasters in the Netherlands. *Prehospital and Disaster Medicine*, 25(2), 127 – 136.
- Legemaate, G. A., Burkle Jr., F. M., & Bierens, J. J. (2012). The evaluation of research methods during disaster exercises: Applicability for improving disaster health management. *Prehosp Disaster Med*, 27(1), 18–26. <http://doi.org/10.1017/S1049023X11006789>
- Lenquist, S. (2005). Education and training in disaster medicine. *Scandinavian Journal of Surgery : SJS : Official Organ for the Finnish Surgical Society and the Scandinavian Surgical Society*, 94(4), 300–310. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16425626>
- Lukosch, H., Kurapati, S., Groen, D., & Verbraeck, A. (2016). Microgames for situated learning: A case study in interdependent planning. *Simulation & Gaming*, 47(3), 346–367.
<http://doi.org/10.1177/1046878116635468>
- Maran, N. J., & Glavin, R. J. (2003). Low- to high-fidelity simulation - a continuum of medical education? *Medical Education*, 37(1), 22–28. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14641635>
- McAulay, L. (2007). Unintended consequences of computer-mediated communications. *Behaviour & Information Technology*, 26(5), 385–398. <http://doi.org/10.1080/01449290500535343>
- McGaghie, W. C., Issenberg, S. B., Petrusa, E. R. and Scalese, R. J. (2006). Effect of practice on standardised learning outcomes in simulation-based medical education. *Medical*

Education, 40, 792–797. <http://doi.org/10.1111/j.1365-2929.2006.02528.x>

- Nilsson, H. (2013). *Demand for Rapid and Accurate Regional Medical Response at Major Incidents*. Linköping University. Retrieved from <http://liu.diva-portal.org/smash/record.jsf?pid=diva2:579248>
- Nilsson, H., Vikström, T., & Rüter, A. (2010). Quality control in disaster medicine training--initial regional medical command and control as an example. *Am J Disaster Med*, 5(1), 35–40. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/20349701>
- Perry, M. (2003). Distributed cognition. In J. M. Carroll (Ed.), *IHCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science* (pp. 212 – 223). San Francisco, USA: Morgan Kauffman Publishers.
- Rybing J., Nilsson H., Jonson C.-O., Bang M. (2016). Studying distributed cognition of simulation-based team training with DiCoT. *Ergonomics*, 59(3), 423-434. <http://doi.org/10.1080/00140139.2015.1074290>
- Rådestad, M., Nilsson, H., Castrén, M., Svensson, L., Rüter, A., & Gryth, D. (2012). Combining performance and outcome indicators can be used in a standardized way for creating lessons learned in disaster response training: An observational study of two of multidisciplinary, full-scale major aircraft exercises. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 20(58). <http://doi.org/10.1186/1757-7241-20-58>
- Sokolowski, J. A., & Banks, C. M. (2010). *Modeling and Simulation Fundamentals: Theretical Underpinnings and Practical Domains*. (J. A. Sokolowski & C. M. Banks, Eds.). Hoboken, NJ: John Wiley & Sons, Inc.
- Sundnes, K., & Birnbaum, M. (2003). Health disaster management: Guidelines for evaluation and research in the Utstein style. Volume I. Conceptual framework of disasters. *Prehosp Disaster Med*, 17(3), 1–177.
- Toups, Z. O., Kerne, A., Hamilton, W. A., & Shahzad, N. (2011). Zero-fidelity simulation of fire emergency response: improving team coordination learning. *Chi 2011*, 1959–1968. <http://doi.org/10.1145/1978942.1979226>
- van Engen-Verheul, M. M., Peute, L. W. P., de Keizer, N. F., Peek, N., & Jaspers, M. W. M. (2016). Optimizing the user interface of a data entry module for an electronic patient record for cardiac rehabilitation: A mixed method usability approach. *International Journal of Medical Informatics*, 87, 15–26. <http://doi.org/10.1016/j.ijmedinf.2015.12.007>