Drivers of evolution of training simulators

Ebba Thora Hvannberg

University of Iceland
Reykjavik
Iceland
ebba@hi.is

ABSTRACT

In many domains, applications have a lifetime of several years, and evolve with new technologies, changing requirements and user behavior. Simulators have been used for training for decades and have evolved in different directions. To learn how applications evolve over time, it is important to understand what stimulates changes. There can be factors in the world external to the application, e.g. human, organizational and technological factors that stimulate changes and there can be factors internal to a software system, such as its structure and complexities. This paper investigates how new work, fidelity and evaluations can contribute to the understanding of evolution of simulators.

CCS CONCEPTS

Software and its engineering~Software evolution
•Human-centered computing~Usability testing.

KEYWORDS

Training simulators; Crisis Management; Fidelity; Evaluation;

1 INTRODUCTION

Not only do systems evolve, but user needs and goals evolve with societal and business changes. Moreover, changing user needs and goals can be attributed to technological developments. Users are known to be rational and adaptable. They change work to adapt to computing misfits, discover workarounds inventively and use software for something that it was not intended for. Conversely, adaptive systems are designed to adjust to user preferences or accommodate user errors. Thus,

changing needs of humans, shifting drivers of businesses and emerging opportunities with innovative technologies encourage watching influential factors of systems' evolution over time.

The topic of system evolution has been addressed by researchers and practitioners of software engineering and Human-Computer Interaction (HCI). Arias et al. [1] have addressed the challenging problem of developing open systems, which provide opportunities for changes to a system during its lifetime, as opposed to closed systems in which all functionality and qualities are fixed when the system is designed. Arias et al. [1] tackled the problem from the user perspective based on several principles: software systems must evolve allowing users to make incremental changes to the core functionality; software systems must evolve at the hands of the users e.g. through end-user programming; and systems must be designed for evolution. Carroll and Rosson [2] described how design rationale is used to create a modified system from a previous one. They state that most technical activities in HCI can be described as transactions between tasks and artifacts, i.e. a task artefact co-evolutionary cycle, suggesting that a new artefact called for a new task, which called for a new artefact and so on. Fischer, Giaccardi, Ye, Sutcliffe and Mehandjiev [3] proposed the SER (Seeding, Evolutionary growth, Reseeding) process model, which encourages designers to describe their designs as meta-designs thus giving users the freedom of being creative designers instead of passive users. Others have built on this work by proposing co-evolution between the technological changes and organizational environments [4].

Training has taken place with the aid of simulators for decades. The origin of training simulators can be traced centuries before they were implemented using Information Technology (IT), such as for military [5] and surgery training [6]. Also, training simulators for air traffic control have a long history [7]. Simulations that are set up to train or study individuals have been termed gaming where gaming involves individuals as decision-makers [8]. Such simulators are used to train a variety of skills including physical skills, knowing facts, forming and following a strategy, tactics and communication.

There are varying triggers for changes and actions in response to these triggers have resulted in revised simulators. To discuss these triggers, we have explored several topics: new work, fidelity and evaluations and

investigated how these topics can contribute to the understanding of evolution of simulators.

2 CRISIS MANAGEMENT TRAINING SIMULATOR

2.1 Introduction

Training for management of airline crisis incidents must be done regularly to practice roles of medical staff, rescue teams, police force and airline personnel. Effectiveness of response to these incidents is paramount and involves training of specialist skills, e.g. medical triaging, providing resources, including medical supplies, skilled personnel and transport, forming strategy and following a plan. Communication between parties and keeping track of resources and casualties are skills that are trained. Training takes place in exercises of various size and form. Small exercises can have five to ten trainees but large-scale exercises can have up to 150 trainees. In a low-fidelity exercise, a table-top exercise takes place around a table training strategic transport of resources, but in a high-fidelity exercise, trainees train on and around an airfield with a simulator airplane and actors playing the role of passengers.

2.2 Emerging incidents and roles

A goal in HCI is to understand practices that are changing. In advocating turn to practice in HCI research, Kuutti and Bannon [9] ask the question how and why these transformations of new practices, often introduced with new IT, take place and how they can be supported. With progress in application areas, new training requirements are bound to emerge. In crisis management, there are changes in response plans, roles, types of resources, e.g. transportation and other equipment. There may be increased variability in the training needs, e.g. between airports [10]. Partly, these changes have been accommodated in training simulators with abstractions of training exercises. Thus, a training simulator tries to anticipate changes in training scenarios through generalization. Training scenarios are elicited from users and set up concretely at the beginning of an exercise. The driver for such abstractions are economical, i.e. it is more expensive to develop software for every new training scenario than to instantiate it from a generalization, and it gives a competitive advantage because large airport service providers want to enjoy the flexibility of training various scenarios. Another mean for economizing has been on using part-task scenarios instead of whole-task scenarios and offering tasks of varying uncertainty [11]. Generation of scenarios in crisis management training with methods of artificial intelligence [12] has been researched.

2.3 Fidelity

Fidelity in the context of simulators was originally defined as the accuracy of the simulator's imitation of the operational equipment, environment and tasks [13]. The concept has developed since and can be viewed from three perspectives: physical, functional and psychological [14]. The success of simulator training relies on how well it manages to give learning feedback to trainees. There are several forms of feedback, originated in the learning context and originated in the system context. Druzhinina, Hvannberg and Halldorsdottir [15] investigated fidelity of feedback by comparing a table-top exercise (TTex) with a real-life exercise (RLex) and a requirements specification of a virtual reality simulator (VRS). Some visual feedback delivery of TTex and VRS was contradictory to the real-life exercise and visual representations differed. There were major deviations between practices of the virtual environment simulator and the real-life exercises. Psychological fidelity could not be found in any of the forms.

Evidence from the aviation industry, military and health professions show that fidelity is difficult to define [16]. Instead of concentrating on fidelity as a goal, it is recommended to focus on methods enhancing transfer of learning. Further, it has been suggested that high-fidelity may not be so important because when trainees see discrepancy between the simulator and the physical model they will suspend disbelief in the simulator for a moment because it may not be that important for the task that they are training for and because they realize that the training is chief [16].

It may prove difficult to follow the fidelity of a moving target. An issue that Mowbray, Holter, Teague and Bybee [17] raised is that increased fidelity of interventions, in our case simulators training, will affect the program model itself, e.g. crisis management. For example, a large component in crisis management training is the post-exercise feedback where it is discussed if training was successful and if the crisis management process needs to be amended.

The discussion on fidelity assumes that the original model is a physical model. Examples in crisis management are vehicles, the airfield, airplanes, passengers or casualties. However, increasingly, the original model is an IT system or a combination of an IT system and a physical model. Examples in crisis management are communication devices such as Tetra, mobile phones and tracking of resources implemented with IT. This requires the simulator to be closely integrated with the crisis management system itself. A simulator's ability to imitate the IT systems can clearly affect the physical and functional fidelity. In air traffic controller's training simulators, we have seen the imitation of the air traffic controllers' IT environment where the simulator is built as an integration with the air traffic controlling environment.

2.4 Evaluations

A systematic evaluation scheme, consisting of a number of benchmarks, exists to assess validity [18]. These benchmarks include face validity, where an expert

evaluates whether the measure assesses what it is supposed to measure, content validity which measures if the content covers what the trainee is to learn, construct validity which assesses if experts show different performance than novices results, concurrent validity where the simulator is compared to other training instruments and predictive validity where it is measured if the performance scores in the simulator predict performance scores in reality.

Wang, DeMaria, Goldberg and Katz [19] reviewed serious games in training of health professionals. Out of a list of 42 serious games, there were 16 training simulations, where situations are realistic and skills can be trained. Out of the 16 training simulators, seven have been evaluated for performance. Of those, four showed improved performance and three showed no significant effect. The three remaining training simulators have been evaluated and showed other factors such as high scores for realism and content, ease of use and usefulness. Studying who participated in the development team, out of 42 serious games, 19 included medical experts, three included trainees on the development team and two included educationalists [19].

Parallel to studying the learning effectiveness of simulator training, it is useful to study their usability and usefulness. Usability evaluation has been conducted of collaboration during crisis management training in a simulator implementing soundscape [20]. A collaborative scenario for two users playing the role of coordinators in the field and in a response center was given. Several method-resources were used to probe for usability problems or mistakes in carrying out crisis management activities. We observed that the functional and the physical fidelity of the simulator was less than expected. We also noted that the training instruments, i.e. the training activities and the roles played were sometimes inadequate. The number of problems we attributed to participants following a script lead us to conclude that one of the biggest challenges in conducting scripted evaluations is to let it not disturb the users and to motivate the users to respond to situations instead of thinking that they can write their own script. We have noticed this when participants take part in real life exercises, and hence it is not unique for virtual training environments. We conclude from these studies that researchers cannot study fidelity of a training simulator in isolation, but need to do so in the context of other training instruments, e.g. the exercises given, the actors and the roles they play.

Given results of evaluations, a question emerges as to what extent outcomes can direct designers, trainees and educators toward more effective designs of simulators and training instruments. This is a topic of a study by Koivisto, Haavisto, Niemi, Haho, Nylund and Multisilta [21] that suggests design principles for simulation games for learning clinical reasoning. One of the learning outcomes of the study is, first, that if students had been involved earlier in the design process, resources could have been saved and, second, that experts from many fields need to be engaged in the design so that the profession being trained for is built in the training simulator.

3 DRIVERS FOR CHANGE

In this section we summarize drivers for change that we have observed from our own research and the literature.

Fidelity: Fidelity has been a driving factor for change in simulators for centuries. Although it is much debated, it is likely that the demand for applying innovative technologies and for more realism may continue to drive that change. Functional fidelity is essential, physical fidelity is argued and psychological fidelity may yet to come. The content between innovation and robustness, which appears as a need to deliver performance, is evident.

New incidents of work and new technologies used as part of work: New incidents of work, be it crisis management, air traffic control or health care, drives training needs. Even training itself can influence work. The use and evolution of technological implementation of work is also a driver for change in training simulators. This may call for integration of training simulators and working systems to a further extent than we have already seen.

Evaluation, effectiveness and economics: The literature has stated the need for robust evaluations and called for validations of training simulators. Much is at stake and there are requirements to validate if training simulators improves trainees' performance beyond other types of training. There is also a requirement to use results of these evaluations as evidence and driver for change in simulator design [22].

The discussion of fidelity is debated with respect to effectiveness and economics. The argument is that low-fidelity is more economical and just as effective as high-fidelity. In areas such as surgery there is much at stake where lack of training can increase patient morbidity and mortality and be financially costly [23]. Development of training simulators with the latest technologies and thorough evaluation at the systems, user and education level is also expensive. With uncertainties in the cost of software development of training simulators and predictive validity it may be difficult to convince buyers to invest in such large systems.

People: Although participatory design and user-centered design have been on the agenda of interactive systems design for decades, there is evidence that training simulators development does not involve trainees. The reason could be that the need for technical skills and innovative technologies may dominate involvement of trainees and their view on practice. A counterargument may be found in a study where operators' participation in design did not respond completely to all the training needs because the technology used in work may require changes in especially roles and responsibilities [24].

Smart education constructs: Although early researchers pointed out the need for a simulator that was not scripted and there was a requirement for a game or a simulator where there was free flow of tasks and collaboration, there has been little coverage of this in the literature except as basic research. Artificial intelligence could be a larger driver for change in the future than it has been in the past, e.g. by allowing simulators to respond more dynamically and over time to trainees' actions and their evolving expertise. A part of this is letting training simulators give trainees feedback and allowing them to reflect on learning. What could help air-traffic control training simulators are advances in simulators that study effectiveness of operations. Examples of such simulators have studied the effectiveness of using artificial intelligence for optimization of air traffic controllers' task load [25].

4 DISCUSSION and CONCLUSION

Although a study such as this cannot be exhaustive, it can help us begin to understand drivers of change of training simulators. For example, artificial intelligence has not been highly visible in this account of training simulators. There were early research efforts in the eighties applying expert systems in training simulators to improve conceptual fidelity [26], but we have not studied if artificial intelligence has been applied in training simulators on the market. The last decade has seen major advances in the area, e.g. agent technology that uses artificial intelligence to coordinate collaborative behavior and can support human and systems behavior in complex operations such as crisis management. However, it remains to be seen applied in training simulators in operation [27]. Although there have been discussions in the HCI literature on evolution and co-evolution, we have not seen detailed studies on change in applications as have been carried out in scientific software [28]. By studying factors likely to have been drivers of change, we may be better prepared to conduct such micro studies. The dynamicity of training simulators and their influence on the need for training and the work to be trained make it a challenging topic of research.

REFERENCES

[1] Arias, E., Eden, H., Fischer, G., Gorman, A. and Scharff, E. Transcending the individual human mind;creating shared understanding through collaborative design. *ACM Trans. Comput.-Hum. Interact.*, 7, 1 (2000), 84-113.

[2] Carroll, J. M. and Rosson, M. B. Deliberated Evolution: Stalking the View Matcher in design space. Human-Computer. *Interaction, 6 (3 and 4), 1992, 281-318.*, 6, 3 and 4 (1992), 281-318.

[3] Fischer, G., Giaccardi, E., Ye, Y., Sutcliffe, A. G. and Mehandjiev, N. Meta-design: a manifesto for end-user development. *Commun. ACM*, 47, 9 (2004), 33-37.

[4] Bourguin, G., Derycke, A. and Tarby, J.-C. *Beyond the Interface: Co-evolution Inside Interactive Systems — A Proposal Founded on Activity Theory.* Springer London, London, 2001.

[5] Smith, R. The Long History of Gaming in Military Training. *Simulation & Gaming*, 41, 1 (2010), 6-19.

[6] Rehder, R., Abd-El-Barr, M., Hooten, K., Weinstock, P., Madsen, J. R. and Cohen, A. R. The role of simulation in neurosurgery. *Child's Nervous System*, 32, 1 (2016), 43-54.

[7] Wickens, C. D., Mavor, A. S. and McGee, J. P. *Flight to the future: Human factors in air traffic control.* National Research Council, National Academies Press, 1997.

[8] Shubik, M. Bibliography on Simulation, Gaming, Artificial Intelligence and Allied Topics. *Journal of the American Statistical Association*, 55, 292 (1960), 736-751.

[9] Kuutti, K. and Bannon, L. J. *The turn to practice in HCI: towards a research agenda.* ACM, Toronto, 2014.

[10] Hvannberg, E. T. and Rudinsky, J. *Crisis Management Training: Techniques for Eliciting and Describing Requirements and Early Designs across Different Incident Types.* Springer Berlin Heidelberg, Orlando, FL, 2011.

[11] Field, J., Rankin, A., Van der Pal, J., Eriksson, H. and Wong, B. W. *Variable uncertainty: scenario design for training adaptive and flexible skills.* Rostock, Germany, 2011.

[12] Grois, E., Hsu, W. H., Voloshin, M. and Wilkins, D. C. *Bayesian network models for generation of crisis management training scenarios.* Madison, Wisconsin, 1998.

[13] Hays, R. T. *Simulator fidelity: A concept paper.* Army research inst. for the behavioral and social sciences, Alexdria, VA. , 1980.

[14] Toups, Z. O., Kerne, A. and Hamilton, W. A. The Team Coordination Game: Zero-fidelity simulation abstracted from fire emergency response practice. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 18, 4 (2011), 23.

[15] Druzhinina, O., Hvannberg, E. T. and Halldorsdottir, G. Feedback Fidelities in Three Different Types of Crisis Management Training Environments. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, 5, 2 (2013), 45-62.

[16] Hamstra, S. J., Brydges, R., Hatala, R., Zendejas, B. and Cook, D. A. Reconsidering Fidelity in Simulation-Based Training. *Academic Medicine*, 89, 3 (2014), 387-392.

[17] Mowbray, C. T., Holter, M. C., Teague, G. B. and Bybee, D. Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation*, 24, 3 (2003/09/01 2003), 315-340.

[18] Gallagher, A., Ritter, E. and Satava, R. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surgical Endoscopy and other Interventional Techniques*, 17, 10 (2003), 1525-1529.

[19] Wang, R., DeMaria, S. J., Goldberg, A. and Katz, D. A Systematic Review of Serious Games in Training Health Care Professionals. *Simulation in Healthcare*, 11, 1 (2016), 41-51.

[20] Hvannberg, E. T., Halldorsdottir, G. and Rudinsky, J. *The Usefulness of Method-Resources for Evaluating a Collaborative Training Simulator.* Springer International Publishing, Bamberg, Germany 2015.

[21] Koivisto, J. M., Haavisto, E., Niemi, H., Haho, P., Nylund, S. and Multisilta, J. Design principles for simulation games for learning clinical reasoning: A design-based research approach. *Nurse Education Today*, 60 (Jan 2018), 114-120.

[22] Cook, D. A. and Hatala, R. Validation of educational assessments: a primer for simulation and beyond. *Advances in Simulation*, 1, 1 (December 07 2016), 31.

[23] Jabbour, N. and Snyderman, C. H. The Economics of Surgical Simulation. *Otolaryngologic Clinics of North America*, 50, 5 (2017), 1029-1036.

[24] Norros, L., Savioja, P. and Koskinen, H. *Core-task design: A Practice-Theory Approach to Human Factors.* Morgan & Claypool, 2015.

[25] Amin, R., Tang, J., Ellejmi, M., Kirby, S. and Abbass, H. A. *Trading-off simulation fidelity and optimization accuracy in air-traffic experiments using differential evolution.* Beijing, China, 2014.

[26] Hollan, J. D., Hutchins, E. L. and Weitzman, L. STEAMER: An interactive inspectable simulation-based training system. *AI magazine*, 5, 2 (1984), 15.

[27] Fiedrich, F. and Burghardt, P. Agent-based systems for disaster management. *Commun. ACM*, 50, 3 (2007), 41-42.

[28] Kelly, D. Determining factors that affect long-term evolution in scientific application software. *Journal of Systems and Software*, 82, 5 (2009), 851-861.