Computers & Industrial Engineering 80 (2015) 171–180

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

A data mining approach for training evaluation in simulation-based training ☆

Juite Wang [a,*], Yung-I Lin [b], Shi-You Hou [a,1]

[a] Graduate Institute of Technology Management, National Chung Hsing University, 250 KuoKuang Road, Taichung 402, Taiwan, ROC
[b] Aeronautical Systems Research Division, National Chung-Shan Institute of Science and Technology, P.O. Box 90008-11-21, Taichung, Taiwan, ROC

ARTICLE INFO

Article history:
Received 25 October 2013
Received in revised form 17 July 2014
Accepted 8 December 2014
Available online 13 December 2014

Keywords:
Training evaluation
Simulation
Multimedia systems
Data mining

ABSTRACT

With the significant evolution of computer technologies, simulation has become a more realistic and effective experiential learning tool to assist in organizational training. Although simulation-based training can improve the effectiveness of training for company employees, there are still many management challenges that need to be overcome. This paper develops a hybrid framework that integrates data mining techniques with the simulation-based training to improve the effectiveness of training evaluation. The concept of confidence-based learning is applied to assess trainees' learning outcomes from the two dimensions of knowledge/skill level and confidence level. Data mining techniques are used to analyze trainees' profiles and data generated from simulation-based training for evaluating trainees' performance and their learning behaviors. The proposed methodology is illustrated with an example of a real case of simulation-based infantry marksmanship training in Taiwan. The results show that the proposed methodology can accurately evaluate trainees' performance and their learning behaviors and can discover latent knowledge for improving trainees' learning outcomes.

1. Introduction

Due to the increasingly complex and changing business environment, enterprise employees not only must possess required professional knowledge and skills, but also need to flexibly adapt their knowledge for use in the changing environment. To develop this adaptive expertise, trainees should be active participants in the learning process and learning should occur in a meaningful or relevant context (Bell & Kozlowski, 2002).

With the significant evolution of computer technologies, simulation has become a more realistic and effective experiential learning tool to assist in organizational training (Bell, Kanar, & Kozlowski, 2008). Simulation is defined as "an artificial environment that is carefully created to manage individuals' experiences of reality" (Bell et al., 2008). Simulation-based training (SBT), therefore, is "the ability to augment, replace, create, and/or manage a learner's actual experience with the world by providing realistic content and embedded instructional features" (Cannon-Bowers & Bowers, 2009). It is highly flexible in terms of place and time of training, which can be used to reduce or eliminate variable costs in traditional training. In addition, SBT can also provide the following advantages (Bell et al., 2008; Cannon-Bowers & Bowers, 2009): safer conditions than real-life situations, minimal influence from external factors, and more opportunities to repeatedly practice rare situations. It has been found that SBT is already used in academic and industrial applications, such as health care (Issenberg, Gordon, Gordon, Safford, & Hart, 2001; Issenberg, McGaghie, Petrusa, Lee, & Scalese, 2005; McGaghie, Issenberg, Petrusa, & Scalese, 2010; Salas, Wilson, Burke, & Priest, 2005), business education (Salas, Wildman, & Piccolo, 2009), pedestrian traffic (Usher & Strawderman, 2010), and disaster prevention (Summerhill et al., 2008).

Although SBT can improve the effectiveness of training for company employees, there are still many management challenges that need to be overcome (Bell et al., 2008; McGaghie et al., 2010). For example, several studies have indicated that the applications of SBT have produced mixed results and have not successfully and effectively grasped the advantages of SBT (Bell et al., 2008; Salas & Cannon-Bowers, 2001). Cannon-Bowers and Bowers (2009) also noted that past simulation-based education efforts have put too much effort on specific technological training systems and too little on training needs. Since SBT has been widely applied in the health care industry, several success factors have been identified in simulation-based medical education (McGaghie et al., 2010). One

☆ This manuscript was processed by Area Editor Paul Savory.

* Corresponding author. Tel.: +886 4 22840547x605; fax: +886 4 22859480.
E-mail addresses: rdwang@nchu.edu.tw (J. Wang), robin.bravo@msa.hinet.net (Yung-I Lin), jo760823@hotmail.com (S.-Y. Hou).
[1] Tel.: +886 4 22840547.

http://dx.doi.org/10.1016/j.cie.2014.12.008
0360-8352/

of the challenges for the development of SBT is how to effectively evaluate training performance and its subsequent impacts (Bell et al., 2008; McGaghie et al., 2010; Salas et al., 2005).

Data mining is the process of exploration and analysis of large quantities of data in order to discover meaningful patterns and rules that can improve business decision making (Berry & Linoff, 2004). It has gradually become an important tool for modern business to transform data into business intelligence and achieve competitive advantage. Accordingly, this paper proposes a hybrid framework that integrates data mining techniques with simulation-based training to improve the effectiveness of training evaluation. The concept of confidence-based learning (CBL) is applied to assess trainees' learning outcomes from the two dimensions of knowledge/skill level and confidence level. Data mining techniques are used to analyze trainees' profiles and data generated from SBT for evaluating trainees' performance and their learning behaviors. The proposed methodology is illustrated with an example of a real case of simulation-based rifle shooting training. From the experimental results, we show that the proposed methodology can accurately evaluate trainees' performance and their learning behaviors and can discover latent knowledge for improving trainees' learning outcomes.

Since SBT usually collects a large amount of data from training sessions, integrating the data mining techniques may be helpful for instructors and trainees to discover useful patterns or rules that can provide immediate feedback for trainees to improve their performance. However, data mining has been rarely used in the field of simulation-based training and this study aims to fill this research gap.

The paper is organized as follows: Section 2 presents the fundamentals of training evaluation and data mining techniques; the proposed methodology is presented in Section 3; the case study and computational experiment are illustrated in Section 4; Section 5 concludes the paper.

2. Fundamentals

2.1. Training evaluation and simulation-based training

Training evaluation is a process that compares the cost of training with the intended learning outcomes assessed in terms of improved performance by trainees (Buckley & Cape, 1990). It can help managers determine if a training program has achieved the desired results and diagnose the strengths and weaknesses of a program for needed improvements. According to Spitzer (1999), training evaluation can turn training into a powerful force for improvement of a business, for both the organization and the people in it.

Several training evaluation models have been developed in the literature (Eseryel, 2002; Holton, Bates, & Ruona, 2000; Moore, Green, & Gallis, 2009). Kirkpatrick's model (Kirkpatrick, 1998) is the most well-known and frequently used model for measuring the effectiveness of training programs in terms of reactions, learning, behavior, and results. The first level measures the immediate reactions of trainees towards training programs (e.g., enjoyment, perceived usefulness, and perceived difficulty). The second level measures the extent to which learning has occurred, where learning is conceived in terms of knowledge, skill, and/or attitude. Further levels measure whether job performance or organizational results have been changed as a result of training (e.g., turnover, volume of activity, cost-cutting, or quality indicators). Regardless of its popularity, Kirkpatrick's model continues to be criticized by researchers for issues, such as liability (Alliger & Janak, 1989) and limited variables and outcome measures (Santos & Stuart, 2003). For example, Phillips (1997, 2003) further developed a framework to compute the return on investment of training. However, challenges still remain in evaluating the effects that the training programs produce in the workspace and in the organization. Since knowledge retention, behavior changes, and organizational impacts resulting from training can only be apparent over time, behavioral and organizational criteria are difficult to measure. However, they are still necessary for training evaluation, because if the desired changes in attitude and behavior do not occur, then the training program is a failure.

SBT has strong potential to create a highly realistic training environment and allow trainees more active participation in the training process. Trainees are expected to act as if they are in a *real* situation. SBT also allows for repeated practice and the quest for excellence through error correction, feedbacks, and debriefing. These help trainees to develop expertise and the necessity for retention of these skills and behavior patterns (Issenberg et al., 2005). Instructors are able to rate trainees' behavior and give feedback to trainees for improving their performance.

Some research has shown that SBT could not only modify trainees' behavioral patterns but also increase their self-efficacy, promoting transfer of training to the workspace (McGaghie et al., 2010). It is well established that self-efficacy enhances learning outcomes and performance (Stevens & Gist, 1997). Similarly, Bruno (1993) also proposed the methodology of confidence-based learning (CBL) which can be used in a learning/training program to measure a trainee's knowledge quality by determining both the correctness of the trainee's knowledge/skill and his/her confidence in that knowledge/skill (see Fig. 1). Once the knowledge/skill correctness and confidence levels have been identified, CBL can identify the learning behavior of a trainee into categories: 'uninformed', 'misinformed', 'doubt', and 'mastery'. Then the instructor can diagnose the learning behavior of a trainee and provide useful feedback to improve the trainee's learning performance. Hunt (2003) also showed that the retention of newly learned knowledge is systematically related to the confidence level people have about the correctness of knowledge. A similar concept was also developed by Jeffries (2005), who included knowledge/skill performance and self-confidence in a simulation-based learning model for nursing. In addition, Yen, Ho, Chen, Chou, and Chen (2010) proposed a confidence-weighting computerized adaptive testing model that provided a more interactive testing environment by focusing on the examinees' confidence in their responses. Their results showed their model yielded ability estimates that were higher and better correlated to examinees' performance in English learning.

2.2. Data mining

The goal of data mining is to extract meaningful patterns and rules from a data set and transform it into an understandable structure for further use (Han & Kamber, 2006; Witten, Frank, & Hall, 2011). Data mining involves various techniques including statistics, neural networks, decision trees, genetic algorithms, and

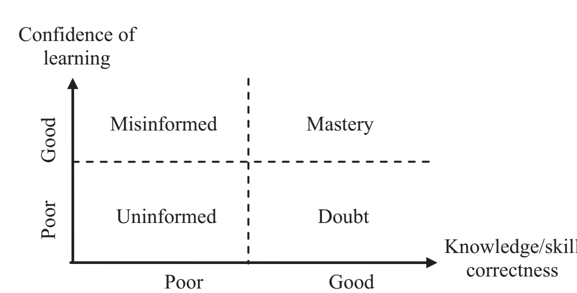

Fig. 1. Classification of learning behavior.