

A Methodology on Guiding Effectiveness-Focused Training of the Weapon Operator Using Big Data and VC Simulations

Won Il Jung, Larry Lowe, Luis Rabelo, and Gene Lee
University of Central Florida

Ojeong Kwon
KAIST

ABSTRACT

Operator training using a weapon in a real-world environment is risky, expensive, time-consuming, and restricted to the given environment. In addition, governments are under intense scrutiny to provide security, yet they must also strive for efficiency and reduce spending. In other words, they must do more with less. Virtual simulation, is usually employed to solve these limitations. As the operator is trained to maximize weapon effectiveness, the effectiveness-focused training can be completed in an economical manner. Unfortunately, the training is completed in limited scenarios without objective levels of training factors for an individual operator to optimize the weapon effectiveness. Thus, the training will not be effective.

For overcoming this problem, we suggest a methodology on guiding effectiveness-focused training of the weapon operator through usability assessments, big data, and Virtual and Constructive (VC) simulations. Usability assessments will be done to identify human factors affecting weapon's effectiveness. Usability assessments consist of data-driven and judgment-driven analytics. The first is generated from raw simulation log files to evaluate the participant's level of skill. The latter is generated from questionnaires, and evaluates the user's background, experience, and opinions about various interactions within the simulation. The human factors must be found out based on several situations to reduce bias. VC simulations are executed under a variety of scenarios, and then, big data are generated such as structured, unstructured, semi-structured types. The big data, including completed questionnaires, are stored in a Hadoop Distributed File System (HDFS) and analyzed with processing via MapReduce, followed by an analytics tool. We can discover important human factors influencing weapon effectiveness, along with optimal values for these factors, which can be a guideline for effectiveness-focused training of the individual weapon operator.

CITATION: Jung, W., Lowe, L., Rabelo, L., Lee, G. et al., "A Methodology on Guiding Effectiveness-Focused Training of the Weapon Operator Using Big Data and VC Simulations," *SAE Int. J. Aerosp.* 10(2):2017, doi:10.4271/2017-01-2018.

INTRODUCTION

In a world with ever-growing threats to safety, coupled with military spending under constant scrutiny, it is critical that governments have the means to evaluate their military power in an accurate, yet cost effective way. In this scenario, Weapon Combat Effectiveness (WCE) plays an essential role in evaluating military power.

Existing literature tends to disregard the variety of factors that influence WCE due to being conducted in limited environments, typically consisting of constructive simulations, minimal weapon systems, and unrealistic custom scenarios. In addition, the subject of human factors, which can be a guideline for effectiveness-focused training of the weapon operator, is ignored completely, due to the lack of virtual simulations. Usability problems often become an overlooked stumbling block, preventing the simulation from fulfilling the training need.

The purpose of our research is to achieve scientific objective levels through WCE analytics, resulting in effectiveness-focused training of weapon system. To achieve this purpose, our goal is to create a unique methodology to identify human factors based on WCE analytics with the aid of big data generated by VC simulations. We will capture the user's interaction in various forms, such as simulation log files, questionnaires, voice, and text messages. Each unique form of data will help answer specific questions about the training that cannot be discovered through the others. The log files will capture entity locations, actions, and timing information, which can be used for analysis and to play back the entire scenario. The questionnaires will aid in capturing data about the human factors of participants. Each piece of this puzzle, which is associated with the operator, is stored in a HDFS. We will use a data-driven approach to assess both the effectiveness of the weapon being operated and the usability of our system through behavioral patterns discovered within the logs.

We will then combine this with a judgment-driven analysis of qualitative data obtained through our surveys. The end result is the objective levels for training that can be validated by Subject Matter Experts (SMEs), and used to aid in decision making for more effective training.

The next section of this paper presents a detailed review of existing literature, which includes WCE and usability assessment methods. Following this, the methodology section covers our approach to find out objective levels for operator human factors to maximize WCE based on big data generated by VC simulations. The remaining sections cover a case study followed by the conclusion and future work.

LITERATURE REVIEW

Weapon Combat Effectiveness

Connors (2015) developed a multi-step agent-based modeling methodology to analyzing weapon effectiveness: 1) define the metrics and collect data, 2) develop simulation scenarios, 3) model verification and validation, 4) specific defined scenarios are analyzed, 5) design of experiment, and 6) analysis of results. These steps were conducted in Lockheed Martin's Cuda prototype, which is a substitute for the Small Advanced Capability Missile (SACM). Several examples were mentioned for methods of effectiveness (MOE) in terms of WCE. In addition, influential factors for each MOE were extracted.

Jiao, Li, Ma, and Yang (2013) created a three-step system for evaluating weapon operational effectiveness: 1) through relation operation, the data for depicting the system character from simulation experiment data is extracted, 2) the value for the evaluation index is gained through complex operations such as arithmetic operation, or functional operation, and 3) operational effectiveness is obtained through the integration of evaluation indexes. The authors praised the flexibility of this approach, particularly in knowledge management. Unfortunately, the analytics were not covered, leaving an inability to visualize this approach in our model.

Jung and Lee (2010) used Army Aviation simulation (AAsim) to estimate optimal requirements of a new attack helicopter flying against anti-armored corps. They incorporated a multi-step methodology: 1) simulation data is identified and entered, 2) the scenario is created, 3) the simulation is generated from the scenario, 4) analytics are used to obtain results, 5) the combat effectiveness is calculated, and 6) the optimal requirement is obtained through estimates. The result was the general methodology for gathering effectiveness analytics, and a calculated number of optimal requirements based on the number of anti-armored corps. Unfortunately, the paper failed to represent a suitable WCE equation for testing under new conditions. In addition, limited data was used, which leads to less accurate results.

Liang et al. (2006) used simulation combined with evolutionary algorithms to evaluate the tactics of using a mixture of decoys and jammers against anti-torpedoes. The model used a torpedo, jammers,

a submarine, and a decoy. This model was based on pre-determined tactics, and the success of submarine evasion was calculated as the MOE. This research is unique because it attempted to optimize effectiveness through evolutionary algorithms.

Lee et al. (2015) looked at the effectiveness of an Unmanned Ground Vehicle (UGV) using ABS, which was modeled along with a tank, and an armored vehicle for three total weapon systems. This was a basic blue versus red team scenario. The blue team had a survival ratio, which was calculated as MOE. They also had a detection and hitting ratio, which were each calculated as their Measures of Performance (MOPs). The model was simulated under differing MOPs within a single scenario. Both MOE and MOPs were compared through sensitivity analysis, to discover that survival rate is tied more heavily to detection rate rather than hitting rate. Thus, an increase in detection rate will have a greater chance of increasing survival rate. This study only used a single scenario, and only compared two MOPs for a single vehicle. This is very limited compared to an actual system, and is not a practical case for calculating realistic MOEs.

Armo (2000) used discrete event simulation to analyze a case study involving a submarine, a counter-measure system, and a light anti-submarine warfare (ASW) torpedo. For the MOE, the probability of the submarine evading the torpedo was calculated. This consisted of varying the submarine speed based on the attack pattern of the torpedo. It was recognized that a submarine's maximum speed is relative its evasiveness within a specified range.

Yoo et al. (2012) created a methodology for evaluating the effectiveness of future weapon systems through One Semi-Automated forces (OneSAF). This system is based on the Product Line Architecture Framework (PLAF), and contains six steps: 1) characteristics analysis of future weapon systems, 2) modeling future weapon systems, 3) implementation of future weapon systems, 4) composition of future weapon systems, 5) simulation, and 6) result analysis. The case study compared a PLAF-based system to a Janus model, and concluded that their PLAF-based system was superior due to convincing results.

Seo et al. (2011) analyzed constructive simulations for an anti-torpedo combat system. This study featured multiple weapon systems, including a torpedo, submarine, surface ship, and a decoy. MOE attributes were compared against the decoy operating system's patterns, operating time, detection range, and speed.

The above literature clearly shows that the human has been vastly ignored in WCE analysis, despite the criticality of the operator's overall experience in effectively using a weapon. In addition, these papers were relegated to using very basic models consisting of a single weapon system, a constructive simulation, and a limited environment. For these reasons, we must find usability measures that affect WCE in robust virtual environments. This data will be used to generate a training profile for the weapon operator, establishing guidelines for more effective training.

Usability Assessment Methods

Data-Driven Analytics

A data-driven approach involves exporting simulation data to log files. The result contains time-stamps, entity attributes, locations, and actions. The raw data is then transformed into useful information through various analytic methods.

Babaian (2007) designed a system for logging user inputs, tasks, and time stamps in an ERP system. Specific events were grouped using a meta-model, and associated with contextual descriptions. The combined information was used to generate usage patterns related to tasks.

Hassenzahl (2000) discusses extracting estimates from participants based on data-driven and judgment-driven methods, and compares the results. The purpose is not to simply discover, but instead to prioritize the usability problems. For the data-driven approach, problem severity is calculated by measuring problem-handling time caused by a usability problem. Meanwhile for judgment-driven, the estimates are gathered through interest groups. It is acknowledged that a data-driven approach is more costly; however it provides an important perspective since the results do not correspond to the judgment-driven estimates. To make an actionable decision in this study, a researcher would need to either decide on a single method, or conduct further research to correlate the results. A key goal of our proposed methodology is to eliminate this roadblock, and automatically correlate data-driven and judgment-driven information.

Judgment-Driven Analytics

Judgment-driven analytics may consist of questionnaires, SME analysis, interviews, focus groups, and a number of other methods. In addition, it may comprise any combination of these data collection strategies. For example, surveys can efficiently gather basic information about a participant's background and experience. This information can be used to find correlation between performance and overlooked traits. A survey can be followed by an interview, or an after action review, to highlight pain-points and possibly answer why a user is struggling. Both sets of data are critical to the research. Some of these data collection mechanisms are more versatile than others, and can be applied before, during, or after the session.

Proctor (2007) tested pilot performance using three different interface configurations for a low-cost virtual flight training simulator. The evaluation protocols involved a background questionnaire, an immersive tendencies questionnaire, a feedback questionnaire, and data analysis comparing the performance of each participant to SME standards. The background questionnaire was distributed first, and obtained general attributes about each participant, such as demographics and experience level relative to pilot training. The second item distributed, an immersive tendencies questionnaire, contained sixteen questions designed to identify the pilot's familiarity and experience with simulated environments. The final item, a feedback questionnaire, provided post-session responses to questions regarding effectiveness of the training, whether or not the system brings any value to the flight school, and overall suggestions for improvement. In addition, the performance data was compared to measures of effectiveness obtained from SMEs and US Army

standards. Some of these performance measures include crash avoidance, the ability to maintain level flight (based on heading, speed, and altitude), and meeting various scenario objectives. This study used a sample of 45 participants to effectively combine performance-based MOEs with a variety of questionnaires to find any statistical significance between human factors and performance on three different simulation platforms. However, this study utilized a limited number of environments, consisting of a practice scenario and a similar final mission scenario. This means that the results could suffer from a bias.

Raybourn (2005) created a multiplayer virtual simulation to train team leaders for the Special Forces. The team obtained user feedback both throughout the development process and after final delivery from eighty-five officers. First, a general background survey was provided, which captured participants gender (all male), age (26-38 years), level of field experience, familiarity with course content, and familiarity with computer games. In addition, another survey concerning general classroom-based simulation usage was distributed to thirty-four Special Forces officers. The results showed that users had an expectation to be engaged, and to gain improved knowledge about personal strengths and weaknesses. This established early requirements for the development team. After delivery, a follow-up questionnaire was circulated to see if initial expectations were met, along with another regarding usability. Fifty-one officers completed these surveys with positive results. In addition, focus groups were also conducted. Unfortunately, the focus group questions and results were not documented. Also, quantitative data was not captured during this phase, so the study was not able to show any significance between human factors and performance.

Silvennoinen (2009) addressed usability challenges with a surgical simulator by using a mixture of methods. A usability expert conducted the heuristic assessment by using design principles from Nielsen (1994) along with several others. Only the most applicable heuristics were used in the evaluation. In addition, a simulator training facilitator was present during this session. Their role was to provide additional support and answer any questions related to operating the simulator. The facilitator's observations were noted along with that of the usability expert. Questionnaires were also distributed to surgical residents to collect their experiences during the training. Lastly, performance data was extracted from simulation parameters to discover performance quality and possible training problems. The usability expert's heuristic analysis uncovered 42 problems with the simulator, mostly concerning aesthetics and visual attributes. Also, nine out of 10 users reported problems in the questionnaires. The results of the three usability tests were combined to find coherence, and to conclude possible training consequences resulting from these issues. Following this, performance data analysis uncovered problems for two out of fifteen users. This combination of measures acts as a fail-safe. First, the team was only able to use two heuristic evaluators, instead of the recommended three by Nielsen. Questionnaires and performance data were used to make up for this deficit. Second, heuristics evaluators may not have a complete understanding of specific end-user requirements. This is why a simulator facilitator was used. This study noted the necessity to

separate problems related to the user interface versus actual skill deficiencies of the participant; however the study failed to actually implement this for existing performance parameters.

This literature survey shows the value of conducting human factors assessments. It is clear that a single method is not sufficient, as most researchers found the most value in using a tiered set of information. The main limitation with existing literature is that their usability assessments were rarely compared to data-driven performance metrics, and were not used to create a learning profile or to accurately assess performance. Instead, they were typically used separately to document the user's overall satisfaction with the simulation. In our approach, we will avoid this by associating the questionnaire results to the raw data in the HDF5. This will be used to generate a learning profile for each participant, and to separate skill-based problems from those related to human factors.

METHODOLOGY

Until present, human factors of operators influencing the weapon system effectiveness were identified and analyzed by using a single weapon simulator under specified scenarios. The existing methods have a critical problem with the accuracy of results. They do not consider the complexity of the weapon system or a variety of scenarios. In addition, important human factors that are identified and analyzed can contain biased results. Others can be missed altogether.

The framework to solve these problems by using big data and VC simulations as shown in Figure 1. It has three steps, which are building model and generating big data are followed by processing big data, and analyzing big data and estimating desired values. This methodology enables us to find out human factors influencing weapon system effectiveness and their desired values.

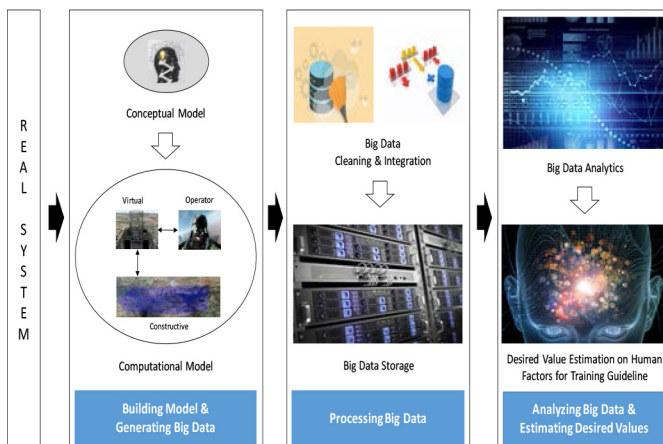


Figure 1. The Framework for Estimating Desired Values of Human Factors to Maximize Weapon System

Building Model & Generating Big Data

This section explains how to build the model using VC simulations and generate big data. First, the conceptual model is developed from real combat systems. The model includes a variety of scenarios under joint battlefield environments, Measures of Effectiveness (MOEs), and human factors to be reflected in the simulation system. The

model is validated by the expert system to be prevented from the inappropriate development. Second, the computational model is formed from the conceptual model. This model is constructed with m constructive simulation models, n virtual simulation models, and a command post as shown in Figure 2. A constructive simulation model consists of a computer-based model for Computer Generated Force (CGF), and a gamer that operates the model only according to the command and control of the command post. A virtual simulation model is composed of a weapon system simulator and its operator(s).

The computational model combines distributed simulation models by using Distributed Interactive Simulation (DIS) and High Level Architecture (HLA) as shown in Figure 3. DIS is used for virtual simulation models, whereas HLA is for constructive simulation models. The combined computational model is implemented according to the conceptual model. Structure, unstructured, and semi-structured data are generated by simulating the computational model. Namely, the big data are generated through interaction between operators and simulators, communication between gamers and a command post, and engagement data between a red and a blue team.

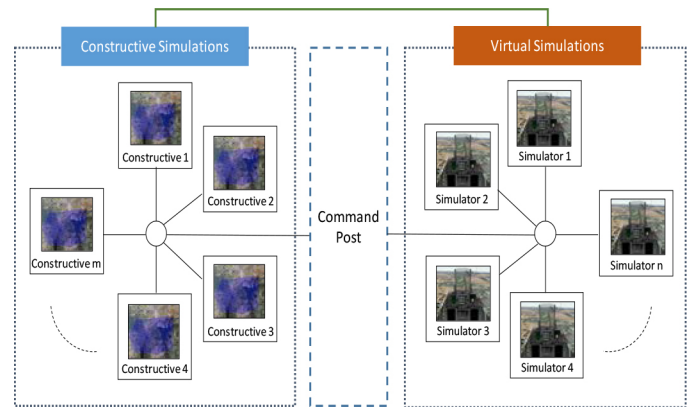


Figure 2. The VC Simulation Model for Generating Big Data

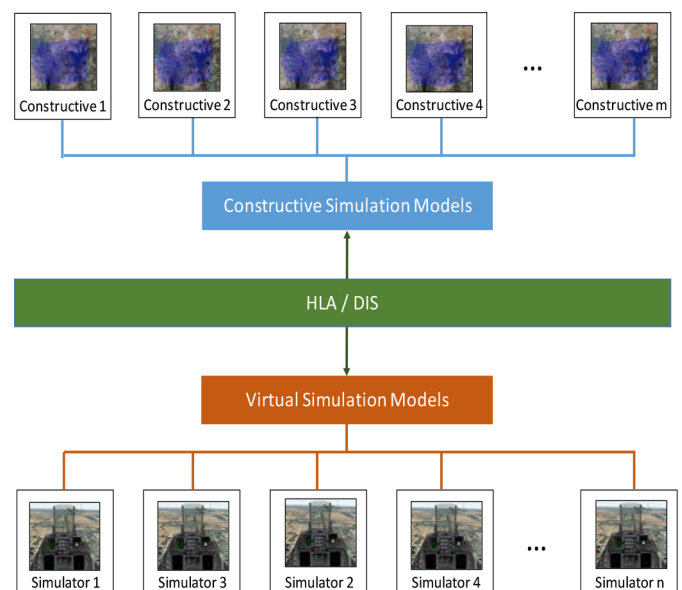


Figure 3. Federation for VC Simulation Models