ARTICLE IN PRESS

International Journal of Industrial Ergonomics xxx (2016) 1–4

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

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

The five key questions of human performance modeling

Changxu Wu

State University of New York at Buffalo, USA

ARTICLE INFO

Article history:
Received 12 July 2015
Received in revised form
12 May 2016
Accepted 14 May 2016
Available online xxx

Keywords:
Human performance modeling

ABSTRACT

Via building computational (typically mathematical and computer simulation) models, human performance modeling (HPM) quantifies, predicts, and maximizes human performance, human-machine system productivity and safety. This paper describes and summarizes the five key questions of human performance modeling: 1) Why we build models of human performance; 2) What the expectations of a good human performance model are; 3) What the procedures and requirements in building and verifying a human performance model are; 4) How we integrate a human performance model with system design; and 5) What the possible future directions of human performance modeling research are. Recent and classic HPM findings are addressed in the five questions to provide new thinking in HPM's motivations, expectations, procedures, system integration and future directions.

1. Introduction

This article elaborates on the five key questions in HPM, describing its motivations, expectations, procedures, system integration and future applications. Since the area of human factors and ergonomics is quite large, this article mainly focuses on the modeling of cognitive-related human performance (e.g., cognition and motor performance under the control of cognition).

2. Q1. Why do we build models of human performance?

In science, besides unifying many scattered findings from empirical studies (Card et al., 1983), models of human performance provide a systematic and computational understanding of the mechanisms of human behavior. In many experimental studies, verbal descriptions (sometimes conceptual models) of mechanisms are very important; however, they cannot make accurate prediction of human performance. Moreover, since the human cognitive and motor system is very complex, verbal descriptions in many cases may not quantify these complex relationships. Computational models of human performance can solve these problems. In addition, models can also guide researchers in data collection and provide researchers with a baseline against which to measure human performance (Sinclair and Drury, 1979).

In engineering, models of human performance can help system designers save significant amount of time and cost in running experiments, and also be integrated into the intelligent/smart systems directly to improve system safety and human performance and/or prevent accidents. The ability to predict human behavior means that, in many cases, accidents are prevented and errors are minimized to improve system safety and efficiency. For example, a human performance model can predict speeding behavior of a driver a few seconds before the actual speeding behavior occurs (Zhao et al., 2013). Once it was embedded in an intelligent system, the system could send pre-speeding warning to drivers to prevent traffic accidents before they occurred (Zhao and Wu, 2013).

3. Q2. What are the expectations of a good human performance model?

The expectations of a good human performance model can be summarized into the following aspects: Mechanisms, Usefulness, Robustness and Generality, and Simplicity (Called as MURGS expectations in HPM).

Mechanisms ("Does this model address the mechanisms of human performance?"): As we discussed in the motivation of human performance modeling, a good model should quantify the relationship between the model's input and output based on the human cognitive and/or motor systems' mechanisms; otherwise, the model may be downgraded to a "black box" model. This issue is related to the *difference between top-down (theory-driven) human performance models and bottom-up (data-driven) models*—including artificial intelligence models (e.g., artificial neural network (ANN) models) and statistical models), since most bottom-up (data

---

E-mail address: seanwu@buffalo.edu.

http://dx.doi.org/10.1016/j.ergon.2016.05.007
0169-8141/© 2016 Published by Elsevier B.V.

Please cite this article in press as: Wu, C., The five key questions of human performance modeling, International Journal of Industrial Ergonomics (2016), http://dx.doi.org/10.1016/j.ergon.2016.05.007

driven) models can relatively easily capture the relationships between model's input and its output (data to be modeled) via model training; however, usually bottom-up (data-driven) models do not quantify the fundamental mechanisms of the human or human-machine systems, or their modeling mechanisms are different from the mechanisms of human cognition and motor system (they have their own sets of modeling/quantification rules). Moreover, due to the lack of the top-down understanding of the mechanisms of human or human-machine systems, bottom-up (data driven) models may over-fit one data set with extensive training for that data set, but under-fit a new data set, leading to their problems in robustness and generality (in other words, leading to the "missing the forest" problem).

Usefulness ("Can this model, once built and verified by the data, improve real-world system performance/safety/efficiency?"): Different from cognitive modeling, such as the work of Isbel and Mahar (2015), that focuses more on the mechanisms and human behavior in lab settings, the emphasis of human performance modeling is more on the human performance and safety in practice and real-world settings. Accordingly, the first expectation of a good human performance model is that its prediction should be directly related to human performance and be useful in real-world system design to improve the performance, safety, and efficiency of human operators and/or the human-machine system as a whole in real-world settings.

Robustness and Generality ("Can this model predict multiple experimental results without over- or under-fitting?"): This expectation of a good model includes two parts: a) Avoid over-fitting or under-fitting; and b) Verification by multiple empirical studies. A good model should not only avoid under-fitting the data (e.g., R square between the model's prediction and experimental data is below 0.5), but also avoid over-fitting the data. Over-fitting usually means a perfect match between the prediction and the experimental data from one study but a poor match between the prediction and the experimental data from another study (See detail discussions of the model over-fitting and under-fitting issues in the work of Lewandowsky and Farrell (2010)). For example, given the same root-mean-square (RMS) of the two models (A and B), Model A verified by two experiments (R square for Experiment $1 = 0.75$ and 0.71 for Experiment 2) is more robust than Model B whose R square for Experiment $1 = 1$ (over-fitting) and 0.46 (under-fitting) for Experiment 2, even if their averaged R square is the same (0.73).

Simplicity ("Is this the simplest model for making a useful and a robust prediction based on human performance mechanisms?"): This expectation is also very important in evaluating a human performance model. This simplicity rule is the same as the parsimonious rule in mathematical and simulation modeling in general: a simpler model is better than a complex model as long as they achieve the same level of functionalities. Moreover, mathematical models are preferred in general than simulation models unless NP-Hard or no analytic solution problem has been encountered by mathematical models (Bank, 2000). The simplicity in HPM is defined as the number of free parameters (The parameters of a model whose values are estimated from the data to be modeled to maximally align the model's prediction) (Lewandowsky and Farrell, 2010), the format and structure of equations if it is a mathematical model, and the number of lines of codes in general if it is a simulation model. For example, a linear model is better than a non-linear model with the same number of parameters as long as both models meet the other three expectations at the same level. Another way to compare the simplicity of different models is to calculate their AIC (Akaike Information Criterion) which considers the number of free parameters (Busemeyer, 2000); however, AIC does not consider structures of equations or the number of lines of computer simulation codes.

4. Q3. What are the procedures and requirements in building and verifying a human performance model?

Depending on the availability of the data and interests of the modeler (the person who builds a model), we summarize the three different approaches to carry out the human performance modeling work.

4.1. Approach 1 (conceptual/existing model or Theory → Model → verify the model by other people later)

In situations that there is no data available or modelers are not able to conduct experiments to verify the model, researchers can still propose/build model without its verification from data. A classic example of modeling work is Einstein's Relativity Theory which was proposed based on theory without experimental data to verify the model's predictions directly at the time when the model was proposed (Einstein, 1905). After a few decades when technologies were feasible to carry out the experiments, the Relativity Theory was eventually verified by the experimental data directly (Hafele and Keating, 1972). We actually think that this is one of acceptable ways of modeling to avoid a modeling problem—If the modeler did have data prior to building a model, he/she could learn what patterns exist in the data during the modeling process and change the model to fit that data, consciously or subconsciously.

In situations that data are available (either from existing published work or from a modeler's own experiments), we typically regard these modeling processes as a mathematical statement proof process (e.g., prove the "$a^2 + b^2 = c^2$" Pythagorean Theorem). This is because the modeler receives data (prediction of the model) before the model is built (although the published modeling work is usually written in reverse order, presenting the model first and model verification with data second). Therefore, a modeler ***should*** clearly provide step-by-step details outlining how his/her model reaches the final prediction (the model's prediction will "definitely" be verified by the data, otherwise the modeler will not even submit this modeling work). If it is a mathematical model of human performance, a modeler should list all of the model derivation steps clearly from the model input to the model output (prediction of the data), without skipping any important steps; If it is a computer simulation model, a modeler should list and ***describe*** the meaning of all the code in the simulation (just listing the computer code at the end of the paper or putting them on a website may not be enough since it is very difficult for a reviewer/reader to understand the code without the author's descriptions), to ensure that there are no codes purposely added in the simulation codes to make the model fit the data.

4.2. Approach 2 (conceptual/existing model or Theory → Model 1 → Verify model (needs Improvement) → Model 2 etc.)

Some modeling work has treated modeling as an iterative process, such as EPIC's modeling study (Kieras and Meyer, 1997). Based on conceptual models, theory, or existing models, researchers built a relatively simple model first, verifying the model with the data; during the verification process, researchers learned that the simpler model is missing some important component(s), and then they improved the model to improve its prediction. The improved model was then verified with the data again.

Please cite this article in press as: Wu, C., The five key questions of human performance modeling, International Journal of Industrial Ergonomics (2016), http://dx.doi.org/10.1016/j.ergon.2016.05.007