The Problem of Manual Assessments of Training Simulators—A Systems Approach

Olatoyosi Olude, Eileen Way, Ph.D.
Department of Systems Science & Industrial Engineering
Binghamton University
State University of New York
Binghamton, NY 13902

Abstract

Since the inception of the Engineering Criteria 2000 (EC2000) by the Accreditation Board for Engineering and Technology (ABET), engineering educators need to demonstrate how well students are learning, which is a challenging task. Many educators have used simulation–based training to assess student learning outcomes. These training simulators have been assessed manually, however, the manual assessments of these training simulators raise issues of reliability and validity. Based on the findings of the case studies used, this paper proposes a systems–based approach, such as Neural Networks can be used when the issue of reliability and validity arises to assess leadership skills of undergraduate engineers.

Introduction

Today the tools and technologies used in the classroom have changed but not the process of teaching. Engineering students all too frequently "sit passively in class copying from the board, reading or just daydreaming while the professor writes on the board"[13].

Recently, engineering educators are changing because corporate businesses are complaining to university administrators about the ability of new graduates to apply their undergraduate education to real world problems. Rugarcia, Felder, Woods and Stice stated that "…corporations are publicly complaining about the lack of professionalism, communication and teamwork etc."[13], in other words, leadership skills.

Implementation of New Educational Programs

There are many solutions colleges find to the question of the lack of leadership in engineers. Engineering educators are exploring many tools and techniques such as class teamwork, field trips, simulation–based training, leadership certificates, licenses etc, some schools have decided to offer various Engineering–MBA programs, or Engineering Management programs. However, these solutions raise a follow–up question on effectiveness. How effective are the above solutions? The question can only be answered through the measurement of an engineer's leadership skill.

Advisors and educators should be involved in assessing leadership skills in undergraduate engineers. Also, improving the quality of an undergraduate education is central to the institution's planning, budgeting and personnel decisions. Two decades ago, Berk referred to Cronbach's principal features of assessment in his research in the 60's, these are[2]:

1. use of a variety of techniques
2. primary reliance on observations
3. integration of information

"Cronbach defined performance assessment as ***the process*** of ***gathering data*** by ***systematic observation*** for ***making decision*** about ***an individual***". He redefined these keywords in the above definition as[2]:

a. ***The process*** instead of a test of any single measurement.

b. ***Gathering data*** using a lot of tools, processes and strategies.

c. ***Systematic observation*** refers to direct observational techniques or now with the use of technology, this can also be referred to as a systems approach to strategic examination.

d. ***Making decision*** is through the data gathered, and it guides the evaluation process.

e. ***An individual*** is either an employee or student, but not a program or product.

Former Methodologies Used for Assessment

Modeling the behavior of training through simulation, also referred to as simulation–based training, is ubiquitous; however, simulation–based assessment can be quite challenging.

What is Simulation?

"Simulation refers to a broad collection of methods and applications to mimic the behavior of real systems, usually on a computer with appropriate software"[9]. Simulation is a generally used term across many disciplines since computers and software are now affordable, involving system designs and models in order to make appropriate decisions.

Why Simulation?

Eggert and Tennyson stated some benefits of using computer simulations which are as follows[5]:

1. Automated computation—Computer simulations automate the model computations.

2. Behavior Animation—Graphical user interfaces animate corresponding changes in position, synchronized and scaled to the numerical analysis time steps.

3. Quick Feedback—The analysis results cannot be overemphasized. Many "what–if" type questions can be quickly explored.

4. Performance Envelope Exploration—Unexpected object or system behavior can surprise and educate users to phenomena that can and does occur in physical models. Users can explore the object's performance envelope, before a physical model is made.

5. Focused Attention—The user's attention is focused when interacting with a computer simulation, such as manipulating the sliders, or inputting variable values. The user cannot be disengaged. The user is actively engaged and learning.

6. Previously Prepared Modules—Valuable student learning time is not wasted on the computer modeling learning curve. Similarly, correct models provide correct results.

Simulation Design and Model

Simulation designs and modelings are problem dependent. For example at Boise State University, Idaho, Eggert and Tennyson introduced system design using computer simulation in the Engineering Mechanics curriculum. The authors discussed the requirements outside of the design arena as follows[5]:

- To simulate a completely defined physical system with the goal of reinforcing basic scientific principles, e.g., conservation of energy in a conservative system.
- For the same physical system simulate with the goal of learning how performance parameters behave in time. For instance, discovering in time the extremes of velocity, acceleration, and force occur and also understanding the behavior of performance parameters of a system.
- Again for the same physical system, varying one design variable at a time in order to discover how the response of various performance parameters changes. As an example, for a system of masses connected by springs, varying a spring constant and studying the change in response of the masses.

Eggert and Tennyson also highlighted the requirements needed within the design arena which are to develop a simulation package, to verify an independent “analysis engine” for use in the evaluation phase of design, to develop a simulation package for the evaluation phase of design and to use a simulation package to verify functional performance of a “design solution” over its range of different operating conditions.

Simulation–Based Training

Simulation–based training tools are assessed by students through surveys and feedback etc. The University of Missouri–Rolla created a simulation recruitment tool for the engineering management department. Nystrom discussed the simulation program that allowed students to experience the real issues of decision making in a high technology company. He also pointed out the importance of business issues such as marketing, management, finance and engineering economics within a team environment that the students experience. As the students allocated resources to design and produce a Palmtop computer, they learned the importance of designing a product line that meets customer's requirements[10].

Nystrom indicated that the simulator was not intended to accurately assess the value of the decisions made by the students, instead the simulation was to assess how much of the simulation helped the students understand the importance of the business related courses taught in the Engineering Management Department[10]. In other words, the student assessments had to be done using alternative methods. At Missouri, students were asked to assess two other sessions from other departments using the ranking below. Nystrom asserted that the strong survey results supported that simulation is an effective way to help students learn the importance of business–related courses[10].

Level Of Understanding	Rankings
Very High	1
High	2
Medium	3
Low	4
Not at all	5

Table 1. Assessment Rankings

Another method of assessment used in academia is using three different techniques to determine effectiveness. An example by Borchert, Jensen and Yates illustrates the transformation from a strictly lecture based education to using a variety of innovative learning techniques. The two innovative teaching tools used are computer based visualizations and hands–on experiments. These tools are used to design an enhanced understanding of specific abstract concepts. Assessment techniques used were with three different tools are stated below[3]:

1. one minute surveys (OMS) taken after each lecture;

2. quick quizzes taken before and after the modules; and

3. specific exam questions designed to measure students' understanding of the concepts covered in the modules

The results of students' assessment of learning to determine effectiveness was based on the correlation between the Myers–Briggs Type Indicator (MBTI) as well as the type of “learner” they are, as measured by the VARK learning style inventory[3]. Borchert, Jensen and Yates found that the hands–on and visual content overall enhances the learning experience.

Furthermore, some authors at the University of Arkansas assessed course performance and ABET outcomes for an industry–based Industrial Engineering senior design course. This was a method used by Rosetti, Cassady and Schneider on an industry–based senior capstone course within the Penn State University Department of Industrial and Manufacturing Engineering. Students were exposed to uncertainties such as change in problem parameters, insufficient data, lack of clarity about the customer's need and corrupt data. Assessment of students' deliverables was based on team accomplishment, peer evaluation, written reports, and an industry sponsor evaluation of team performance[12].

All these training simulators and assessments are in the interest of teaching engineers great leadership skills. Rochefort noted that educators often ask Industrial Advisory Board (IAB) which qualities and skills they would most like to see in new engineering graduates. He claimed that good technical skills in the given discipline are always a priority and are closely monitored with the grading systems. However, good oral and written communication skills, the ability to work in interdisciplinary teams, and leadership skills are very important[11]. He stated that all evaluations are maintained confidential by the instructor to encourage honest and direct assessments by students and mentors.

All these examples show that teaching and training are often simulation–based, whereas the assessments of all these simulators are manual. Manual assessments to these simulators raise issues of reliability and validity.

Simulation–Based Assessment

Streufert, Pogash and Piasecki found out that researchers employed constructs and measurement techniques of complexity theory to test performance in complex perceptual and decision–making tasks. They also found out that cognitively complex individuals differ from those with lesser complexity in terms of content and flexibility of attitudes. In simulation–based assessment, "…a person's cognitive complexity—how to think and behave might need to be modeled in order to make appropriate decision–making"[14], which is an extremely daunting task because it is difficult for people to express their thought processes for making decisions.

Streufert et al. pointed out that people are generally quite aware of what they think or do but find it challenging to conceptualize how they thought about it and made decisions. So responses to an assessment should reflect the thoughts of the participant, however if the examiner is unable to conceptualize the decision–making process of the participant, then the responses might be irrelevant.

Another issue of Streufert et al. is bias. Since there are different styles of decision–making and information processing, these styles might affect competence and success. They proposed that assessment should not be biased by task–specific knowledge, experience or training.

Streufert et al. designed a simulator for assessing managerial competence where two scenarios were presented in random order to the participants. Performance scores for each participant were calculated and separated into three groups which are[14]:

1. Measures of content (i.e. what the decision maker did)
2. Measures of structural style (how the decision maker approached the problems at hand)
3. Mixed measures

The measures of managerial style used by Streufert et al. focused on twenty different attributes and some of these are: Diversity of Action, Use of strategy, Systematic Functioning, Systematic Approach to Strategic Planning, Length of Forward Planning.

The simulation technology measured the reliability values of the simulation to be highly inter–correlated. They found that two simulations generated reliable data and reliability was demonstrated during the measurement of planning and strategic action. However, the issue of subjectivity was not addressed.

In summary, the use of simulation–based training, teaching and assessment raises the issues of reliability and validity such as bias, subjectivity, etc.

System–Based Approach

Assessment centers manually assess training to employ multiple measurement techniques that focus on content–specific knowledge, experience etc. while the simulation technology employs single lengthy and realistic task which generate multiple performance measures and analysis that are attained through a computer program.

On the other hand, system–based methodologies such as Fuzzy Logic[8], Neural Network (NN)[6] and other hybrids[1] can also produce solutions to problems that involve reliability and validity issues. There are numerous types of neural networks[6,7]. One type of neural network that can be used to assess leadership skills in undergraduate students is Backpropagation Network (BPN).

Backpropagation Network

A standard Backpropagation network (BPN) also known as, Multi–Layer Perceptron (MLP) is a fully connected feed–forward network i.e. all connected links are adjacent and the input signals flow from left to right. Backpropagation network undergoes a supervised learning process, and the output signal goes through a binary sigmoid activation function. Training backpropagation involves three stages: the feedforward of the input training pattern, the calculation and backpropagation of the associated error, and the adjustment of the weights. The aim of training backpropagation is to achieve a balance between the ability to respond correctly to the input patterns that are used for training and the ability to give reasonable responses to input that is similar to that used in training[6].

Case Study

Leadership skills, such as communication and teamwork, can be assessed through the courses (i.e. number of hours spent in the classroom on a leadership course) taken by each student using a student's minor Grade Point Average (GPA).

Backpropagation network is used to model this system and in this case, leadership assessment begins with building a network of two inputs, communication (x_1) and teamwork (x_2) which are pre–processed. The inputs send signals to three hidden nodes which backpropagate the error through training of the network, and eventually sends the signals to the output, which is post–processed. Using the algorithm to the Backpropagation Network below, the network architecture is also illustrated in Figure 1.

1. `Initialize the weights with random values`
2. `Propagate first training vector`
3. `Calculate output`
4. `Calculate error`
5. `Propagate error signal & calculate weight changes`
6. `Adjust weights`
7. `Go back to #2, loop until stopping criteria is met`

Finally, the output is computed for each student's GPA and the root mean square error (RMSE) is calculated. The lower the RMSE the better the prediction. In this case, backpropagation network has a RMSE training = 0.4388 and a RMSE testing = 0.3910.

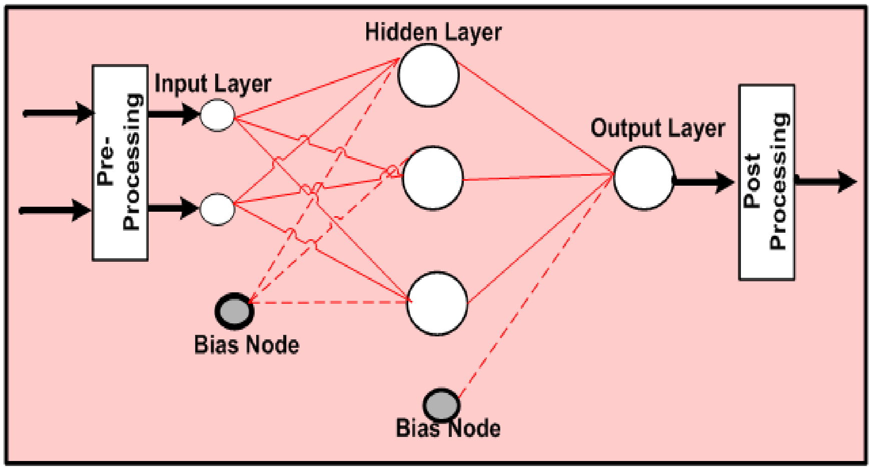

Figure 1. Backpropagation Network

Reliability & Validity

Backpropagation has an ability to generalize through learning. It makes generalizations by "producing reasonable outputs for inputs not encountered during training or learning"[7]. Backpropagation network is capable of adapting to its surrounding environment[7]. This is done through training of the network. Since backpropagation network is fault–tolerant[15] i.e. the property that enables a system to continue operating properly in the event of the failure of some of its components. A Backpropagation network is reliable if the root mean square error of the testing patterns is closer to 0.

A backpropagation recognizes inherent vector features present in the data set and classifies them accordingly. This feature is enhanced through the universal approximator property defined by Kolmogorov's existence theorem which states as follows: "A feed–forward Neural Network with three layers of neurons (input units, hidden units, and output units) can represent any continuous function exactly"[6].

However, backpropagation is not an ideal model, so there exist some variance and bias which should be minimized. Statistical literature contains a variety of error estimation methods (e.g. resubstitution, cross validation (CV), jackknife, bootstrap and train–and–test)[4,16]. Efron and Tibshirani, Twomey and Smith concluded that backpropagation network (BPN) is effectively valid, if the data is valid and the train–and–test method is used[4,16].

Twomey and Smith also investigated the neural network standard that should be adopted when faced with constrained amounts of data. They also concluded that the traditional train–and–test validation "using nearly all the sample for training (e.g. 90%) is highly dependent on the small number of observations left in the test set"[16].

Conclusion

Measuring the leadership skills of undergraduate engineers should be implemented using the backpropagation network because backpropagation is reliable and can deal with the issue of validity. In addition, backpropagation network learns leadership patterns from the training data.

However, backpropagation network is limited to dealing with subjectivity. The issue of subjectivity raises the need to define leadership in engineers. Other system–based approaches such as fuzzy logic techniques and hybrids such as neuro–fuzzy systems are under investigation. This research leads to future directions in using techniques such as text mining and fuzzy inference system to define leadership in engineers.

Bibliography

[1] Afolabi, M. O. and Olude, O. 2007, Predicting Stock Prices Using a Hybrid Kohonen Self Organizing Map (SOM), *in* 'Proceedings of the 40th Annual Hawaii International Conference On System Sciences (CD-ROM)', IEEE Computer Society Press, Hawaii, pp. 1–8. (In Press).

[2] Berk, R. A. 1986, *Performance Assessment: Methods and Applications*, John Hopkins University Press, MD.

[3] Borchert, R., Jensen, D. and Yates, D. 1999, Hands-on and Visualization Modules for Enhancement of Learning in Mechanics: Development and Assessment in the Context of Myers Briggs Types and VARK Learning Styles, *in* 'Proceedings of the 1999 ASEE Annual Conference and Exposition', Charlotte, NC.

[4] Efron, B. and Tibshirani, R. 1986, 'Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy', *Stat. Sci.* **1**, 54–77.

[5] Eggert, R. J. and Tennyson, S. 1998, Re-Engineering Open-Ended Problems and Computer Simulation for Effective Development of Student Design Skills, *in* 'Proceedings of the 1998 ASEE Annual Conference and Exposition', Seattle, WA.

[6] Fausett, L. 1994, *Fundamentals of Neural Networks: Architecture, Algorithms and Applications*, Prentice Hall, MA.

[7] Haykin, S. 1999, *Neural Network: A Comprehensive Foundation*, Prentice Hall, MA.

[8] Jang, J., Sun, C. and Mizutani, E. 1997, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, MA.

[9] Kelton, W., Sadowski, R. P. and Sadowski, D. A. 2002, *Simulation with Arena*, McGraw Hill, NY.

[10] Nystrom, H. E. 1999, A Simulation Recruitment Tool for Engineering Management, *in* 'Proceedings of the 1999 ASEE Annual Conference and Exposition', Charlotte, NC.

[11] Rochefort, W. E. 1998, Leadership and Mentoring Undergraduate Engineering Programs, *in* 'Proceedings of the 1998 ASEE Annual Conference and Exposition', Seattle, WA.

[12] Rosetti, M. D., Cassady, C. R. and Schneider, K. 2002, Assessing an Industry-Based IE Senior Design Course, *in* 'Proceedings of the 2002 ASEE Annual Conference and Exposition', Montreal, Canada.

[13] Rugarcia, A., Felder, R., Woods, D. and Stice, J. 2000, 'The Future of Engineering Education. Part 1: A Vision for a New Century', *Chemical Engineering Education* **34**(1), 16–25.

[14] Streufert, S., Pogash, R. and Piasecki, M. 1988, 'Simulation-Based Assessment of Managerial Competence: Reliability and Validity', *Personnel Psychology* **41**(3), 537–557.

[15] Tsoukalas, L. and Uhrig, R. 1997, *Fuzzy and Neural Approaches in Engineering*, Wiley and Sons.

[16] Twomey, J. M. and Smith, A. E. 1998, 'Bias and Variance of Validity Methods for Function Approximation Neural Networks Under Conditions of Sparse Data.', *IEEE Transactions on Systems, Man and Cybernetics—Part C: Applications and Reviews* **28**(3), 417–430.

Acknowledgement

The authors thank the support of Clifford D. Clark Fellowship that enabled them to complete this project. They also appreciate the support of the Alliance for Graduate Education and the Professoriate (AGEP) funded by National Science Foundation (NSF).

Biography

OLATOYOSI (OLA) OLUDE is a Ph.D. candidate in Systems Science at the T. J. Watson School of Engineering and Applied Science, SUNY, Binghamton. Her research interests include Systems Engineering Analysis & Design Principles, Systems–based methods to Leadership Assessment in Engineers and Systems Methodologies such as Soft Computing, Fuzzy Logic, Neural Network and hybrids. She has co–authored a conference proceeding for IEEE Computer Society Press entitled *Predicting Stock Prices Using a Hybrid Kohonen Self Organizing Map (SOM)*, presently *in press*. She has over ten years of IT experience relating to business, technical and research functions. Ola is a member of IEEE, National Society of Professional Engineers and *presently* the Public Relations of the American Society for Engineering Education–St. Lawrence Section.

EILEEN C. WAY, PH.D. is an Associate Professor of Systems Science & Industrial Engineering at the Thomas. J. Watson School of Engineering and Applied Science, SUNY, Binghamton. Dr. Way is the author of *Knowledge Representation and Metaphor* and (with J. Aronson and R. Harre) *Realism Rescued: How Scientific Progress is Possible*. Journal papers she has published deal with subjects such as the role of similarity in models, connectionism and conceptual structure, the significance of AI for Psychology, artificial intelligence approaches to models and metaphor and the cognitive content of metaphor. Dr. Way teaches introduction to systems science, computability and logic and has been teaching the undergraduate Systems Design and Capstone Design course sequence. She is a member of the Cognitive Science Society, the Society for Machines and Mentality, and the American Society for Engineering Education.