Scheduling Training to Manage Acquisition & Decay

Mohammed Eslami, Ph.D.; Jared Freeman, Ph.D.
Aptima, Inc.
Washington, DC
meslami@aptima.com, freeman@aptima.com

Scott Pappada, Ph.D.
Aptima, Inc.
Dayton, OH
spappada@aptima.com

ABSTRACT

The accelerating effects of adaptive training systems are well established (Lesgold, 2012; Cohn & Fletcher, 2010). This power might be enhanced further by scheduling training to accelerate acquisition and scheduling re-training to reduce decay. Models of acquisition and decay have been available to support scheduling since the early days of memory research (Ebbinghaus, 1913). But these models are derived mainly from laboratory tasks that are learned and executed over seconds or minutes, and performed in isolation from competing tasks. The models are much less explanatory or predictive over real world tasks that are complex, learned and executed over hours and days, and situated in a river of daily assignments that impose the scientifically acknowledged cause of skill decay: interference with memory retrieval (Farr, 1987; Arthur, et al., 1998). In this paper, we propose a new approach to acquisition and decay modeling to make the science of skill acquisition and decay more useful and usable. The approach applies machine learning techniques to model skill acquisition and decay. We apply these methods to a large dataset from a game-ified working memory exercise, compare the performance of these methods with a conventional technique, and present the argument for applying these methods to predict learning and schedule training for realistically complex tasks such as system diagnosis and corrective maintenance.

ABOUT THE AUTHORS

Mohammed Eslami, Ph.D. is a Senior Research Engineer with Aptima, Inc. His background lies in smart surveillance systems, statistical signal processing and computer vision. His interests include using mathematical models for multi-objective constraint optimization, intelligent tutoring systems, data fusion, and big data analytics. Dr. Eslami received his Ph.D. in Electrical Engineering and a B.S. in Electrical Engineering and Mathematics from the University of Maryland, College Park (UMCP). Prior to joining Aptima, he was a post-doctoral fellow at the Food and Drug Administration and currently holds an Associate Research Faculty position at the UMCP Department of Electrical and Computer Engineering working on the development of the next generation traffic surveillance system for the Federal Highway Administration.

Jared Freeman, Ph.D., is Chief Scientist of Aptima Inc. He investigates problem solving by individuals and teams in real-world settings. Dr. Freeman develops decision aids, training content and systems, performance and communications measures, feedback technologies, and organizational structures that support mission leaders and their staffs. Dr. Freeman is the author of more than 120 articles in journals, proceedings, and books concerning these and related topics. He holds a Ph.D. in Human Learning and Cognition from Columbia University and a M.A. in Educational Technology from Teachers College, Columbia University. Dr. Freeman is a Contributing Editor to the journal Human Factors.

Scott Pappada, Ph.D., is a Senior Biomedical Engineer with Aptima, Inc. He has a broad background in biomedical engineering specifically focused in the field of advanced signal processing and signal modeling of various physiologic signals. Particular areas of expertise include pattern recognition algorithms, and machine learning (i.e. artificial neural networks). Dr. Pappada's research involves development of medical device algorithms, technologies, and software which facilitate intelligent therapeutic direction and clinical decision support from both clinician and patient perspectives. Dr. Pappada received his Ph.D. in Engineering with a Bioengineering concentration from the University of Toledo, and a Bachelors of Science in Biomedical Engineering with a minor in Biomedical Sciences from Marquette University.

Scheduling Training to Manage Acquisition & Decay

Mohammed Eslami, Ph.D.; Jared Freeman, Ph.D.
Aptima, Inc.
Washington, D.C
meslami@aptima.com, freeman@aptima.com

Scott Pappada, Ph.D.
Aptima, Inc.
Dayton, OH
spappada@aptima.com

INTRODUCTION

Training is efficient when it accelerates the acquisition of skills and slows their decay. This entails measuring learning, modeling acquisition and decay, and using such models to specify when and what the trainee can most profitably study to build skills or retain them. Few if any training systems do this, unfortunately. A robust measurement and modeling solution that forecasts and manages learning in training and on the job would enable the military to implement training that is more efficient, and talent management that delivers and applies trained personnel more effectively.

To realize such systems requires new modeling techniques. Models of acquisition and decay have been available to support scheduling from the early days of memory research (Ebbinghaus, 1913) through today (Fitts and Posner, 1967; Pavlik and Anderson, 2003; Jastrzembski, et al., 2006; Wang, et al., 2013). But these models are derived mainly from laboratory tasks that are learned and executed over seconds or minutes, and performed in isolation. The models are much less explanatory or predictive over real world tasks that are complex, learned and executed over hours and days, and situated in a river of daily assignments that impose the scientifically acknowledged cause of skill decay: interference with memory retrieval (Farr, 1987; Arthur, et al., 1998).

In this paper, we explore methods of analyzing skill acquisition and decay using the traditional General Performance Equation (GPE) and two machine learning models: Hidden Markov Models (HMM) and Multi-Layer Perceptron Models (MLP). As we report, below, the alternatives to the GPE produced respectable accuracy scores: Mean Absolute Deviation (MAD, a measure of error) was 8% for the GPE, 10% for the HMM, and 27.5% and 18.0% for iterative and fixed trajectory MLP models respectively. Importantly, these models offer functionality that the GPE cannot. They can:

1. Automatically refine their estimates of acquisition and decay over time, as data concerning student performance accumulates. This enables the system to model skill acquisition and decay curves roughly at first, by drawing on the research literature, and to gradually develop curves that fit real-world training and operational phenomena well.
2. Incorporate a relatively rich array of factors that influence learning, such as characteristics of students (e.g., their experience with related tasks, their physiological state, age), knowledge and skill requirements, problems or tasks, and training environments.
3. Represent learning and acquisition curves, either in whole or in part, that do not have the canonical form determined by the GPE.
4. Discriminate between specific students or sections of the student population that exhibit different learning patterns, thus improving the fit of predictions and the quality of recommendations to the student(s).
5. Prescribe an optimal training schedule, by enabling an instructor or a Monte Carlo curriculum generator to test the effect of alternative schedules on acquisition and decay.

Applications

Systems for scheduling training have high value, directly and indirectly, to Defense operations. Consider the talent management challenge of building and sustaining a corps of Naval system maintenance technicians. Navy systems are complex and often vary in their form and functionality between vessels. Navy maintenance personnel have sound training, but informal reports by operators suggest that diagnostic expertise is hard won in training, and hard to maintain in the operational environment.

Few current training systems integrate performance assessment, fewer still adapt instruction to performance, and none to our knowledge adapt instructional schedules to optimize learning curves. An efficient, effective training technology could accelerate the acquisition of maintenance knowledge and skill, and maintain those skills over time to the extent that it reliably recommends the right (re)training at the right time to technicians, their instructors, or their supervisors. The recommender component requires data concerning the activities and performance of maintainers in training and on the job. Objective, reliable data might be gathered in two ways. We might instrument technicians and technology with sensors that log maintenance training and job activity to support inferences about proficiency. A more modest approach would embed activity and proficiency assessment into training systems only (as illustrated in Figure 1). Either approach would provide data needed for a robust adaptive training recommender. Either could undergird a quantified talent management system that enables the DoD to assess and restructure the pipeline from accession through training, deployment, and retirement. It would enable DoD to assess the effectiveness of training regimens, job assignments, and career paths. Finally, it would inform research concerning all of these matters.

Figure 1: Flow of a training system that embeds activity and proficiency assessment into the tool itself to assess learning of complex tasks and adapt instructional schedules to learning curve forecasts.

In the remainder of this paper, we set aside the challenge of measuring proficiency and turn to the problem of predicting the acquisition and decay of skills using from good measurement data.

MODELS FOR SCHEDULING

Current methods of computing knowledge acquisition and decay curves interpolate points that are collected based on successful performance of a single, specific training task. These interpolation techniques assume fixed, specific forms of learning behavior among students, and they do not represent the range of factors that may influence learning by a specific population of students in a specific domain under specific learning conditions. Skill acquisition and decay rates for an individual or group can be modeled as a function of the numerous data sources collected before, during, and following a training session. In this sense, prediction of skill acquisition and learning curves becomes a data-driven problem, one that requires the capabilities, enumerated above, of HMMs, MLPs, and other machine learning models.

Below, we describe and test three models of acquisition and decay. We begin with the standard regression model, which serves as a benchmark: the General Performance Equation; we then investigate Hidden Markov Models (HMM), and Multi-layer Perceptrons (MLP) as the two machine learning models for comparison. To our knowledge this is the first comparison of these machine learning models to the standard GPE.

The General Performance Equation (GPE)

The General Performance Equation (GPE; Anderson and Schunn, 2000) specifies the power law of practice, the power law of forgetting, and the multiplicative effect of practice and retention. Specifically, it is given by:

$$GPE = AN^c\,T^{-d} \qquad (1)$$

where *A* is a free parameter scalar, *N* is the amount of practice, *c* is the rate of learning, *T* is the time since learning, and *d* represents memory decay. Parameters *A*, *c*, and *d* are found to best fit the data (Jastrzembski, Gluck, and Gunzelman, 2006). These model parameters are now fixed for a student and implicitly assume that the student will not deviate from these parameters as their mastery grows or shrinks. However, since this equation is widely viewed as the standard in the literature and has been shown to work well with a fixed set of data, we will use it as our ground truth to compare the performance of the machine learning models described below.

Hidden Markov Models (HMMs)

A Hidden Markov Model (HMM) is describes the transitions between a trainee's states (e.g., Novice (N), Intermediate (I), Expert (E)) given a set of measures of performance or observations (O). In Figure 2, circles represent the state of a student, while O|* represents the probability of an observation given a student is in a certain state: N, I, or E. Specifically, with known past observations of a student, the Baum-Welch algorithm is used to infer the transitions of a student between each state, as well as the probability of emitting an observation from a specific state. As an example, a new electronics technician is most likely to start off in the state novice. When in the state novice, their performance on test items requiring expert K&S will likely be between 0-33%. As they advance in training, they will likely move to the I state where their performance will be between 34%-66%, and lastly they will advance to being an expert with performance between 67%-100%.

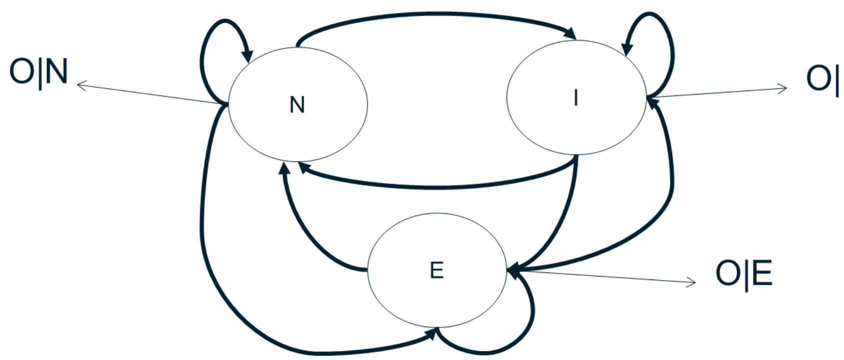

Figure 2: A HMM representing learning and decay. The nodes represent skill levels (N=Novice, I=Intermediate, E=Expert) on a task, and the links between the nodes represent the probability of transition from one level to the next. The O|* are the observations that are made from the training cycle shown in Figure 1 to infer the level of a trainee.

Multi-layer Perceptrons (MLP)

As a final model for comparison Multi-Layer Perceptron-based Models were developed to predict the learning curve of a learner/trainee over *N* future sessions. Multi-Layer Perceptrons (MLPs) are a common type of feed forward artificial neural network. For this effort, we investigated the development of two different types of MLP-based models: iterative MLPs and fixed trajectory MLPs. An iterative MLP predicts the learner/trainee performance for the next session (i.e. one session in the future) and then use its output iteratively to generate an estimate of a learning curve over N future sessions. Conversely, fixed trajectory MLPs are similar in design and structure to the iterative MLP with the only difference being that model outputs are not used iteratively (i.e., there are fixed model inputs based on concurrent and historical available input data) and the model is constrained to generate a fixed number of outputs *N* (as an estimate of a learning curve over *N* future sessions). The structure of an iterative MLP implemented in this effort is shown below in Figure 3. MLPs and other neural network-based models have a structure consisting of three types of layers: 1) an input layer, 2) hidden processing layer, and 3) an output layer. All MLPs developed in this effort were trained via back-propagation and used the Levenberg-Marquardt algorithm to determine optimal model weights. (Back-propagation is short for "backward propagation of errors" and is a common method of training neural network models.) During model training, the error of the model with respect to a desired/target output is "back-propagated" to the various layers of the neural network model, where weights are continuously adapted via an optimization method (Levenberg-Marquardt algorithm in this case) in order to identify a set of weights which yield optimal model performance (i.e. minimization of model error).

The input layer of the MLP consists of N processing elements based on the number of features (i.e. inputs) used by the model. For the MLPs developed in this study, model inputs included: performance of a learner/trainee (on a 0-100 scale), the intended difficulty of a lesson/task, the amount of time between training sessions (in days), and the duration of a training session/lesson (in minutes). These inputs were included in the model for the most recent session and N historical sessions. The weights of each perceptron, w_{ij}, were found by using the Levenberg-Marquardt optimization algorithm. The advantage it provides over the GPE is, again, that the weights are changed dynamically to better fit and predict the learning curve.

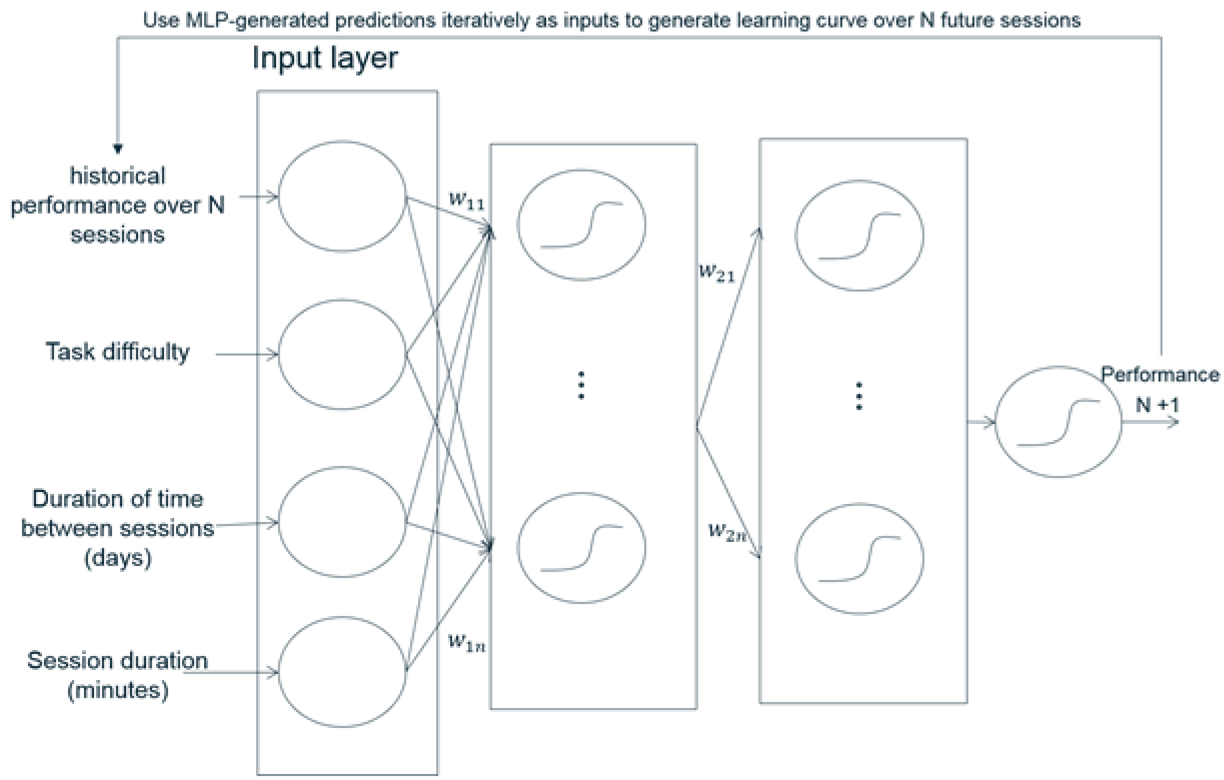

Figure 3: An iterative MLP-based model for prediction of learner/trainee learning curves takes an easily extensible set of input measures and identifies the appropriate weights to model the effect each measure has on the performance of a trainee. Layers of sigmoid functions are used to transform the input measures to the appropriate performance score.

The architecture of this MLP represents fairly few and conventional learning factors. However, the MLP is most appropriate when we must represent a large variety of features (e.g., time delays, difficulty, student characteristics, task characteristics), or when the history of training events strongly influences learning.

EXPERIMENTAL TRIALS & RESULTS

The three models described have different input and output parameters, making a direct comparison challenging. Specifically, the GPE's inputs are the trial number and the spacing between sessions, and its output is the accuracy, given a time delay between training sessions. The HMM takes as input the sequence of accuracy observations, and outputs a new predicted sequence of accuracies at each trial. Lastly, the MLP, the most generalized time-series model, takes multiple features as input, such as difficulty, time delay, and session duration, to output a predicted time-series of the accuracies.

To compare the predictions of each model with actual learning data, we used the mean absolute difference for N training points as a figure of merit. Mean Absolute Deviation (MAD) was computed thus:

$$MAD = \frac{\sum_{i=1}^{N} |ActualPerf(t_i) - PredictedPerf(t_i)|}{N} * 100 \qquad (2)$$

Dataset: Cognitive Skills Training

We tested these models on data that documents the acquisition and decay of well-defined and fundamental cognitive skills in a Western-themed game called Mind Frontiers (Ward, Paul, et. al, in review), by 241 anonymized participants over a 17 week period of play. Participants played 7 games at 9 different difficulty levels. Each game was specifically designed to exercise one of the following functions: working memory, reasoning, switching (or categorization), or executive function. The participants executed more than 815,000 trials (or "turns") of one of these games: an implementation of the dual n-back task (Jaeggi, et al., 2008). Each trial lasted three seconds. Blocks of trials were separated by one or more days during which skill decay could theoretically occur. Thus, this dataset was thus well-suited to the planned analysis: it was collected under experimental conditions that ensured the quality of the data, it could contain evidence of both acquisition and decay of a skill over an interestingly long time span of months, and it was large enough (both in terms of participants and trials) to support the development of new models using certain machine learning applications.

The General Performance Equation (GPE)

The GPE is a simple regression model that fits a power-law curve of positive exponent for acquisition, and negative exponent for decay to the data. In our dataset, each user contributed multiple trials from about 109 sessions. These sessions were separated by a variable time difference between consecutive sessions. That is, within these 109 instances, there were multiple instances where there was a gap of, for example, 1 day, and other instances where there were gaps of 30 minutes. With these considerations in mind, the data was processed thus for the GPE. Focusing on a specific user session at a specified difficulty level, the time deltas between consecutive session start times were found. The accuracy at a specific trial was the average accuracy between consecutive trials to correspond with the appropriate time delta between trials. The results for one specific participant chosen at random from the dataset are shown in Figure 4 which achieved an error of 6.3%.

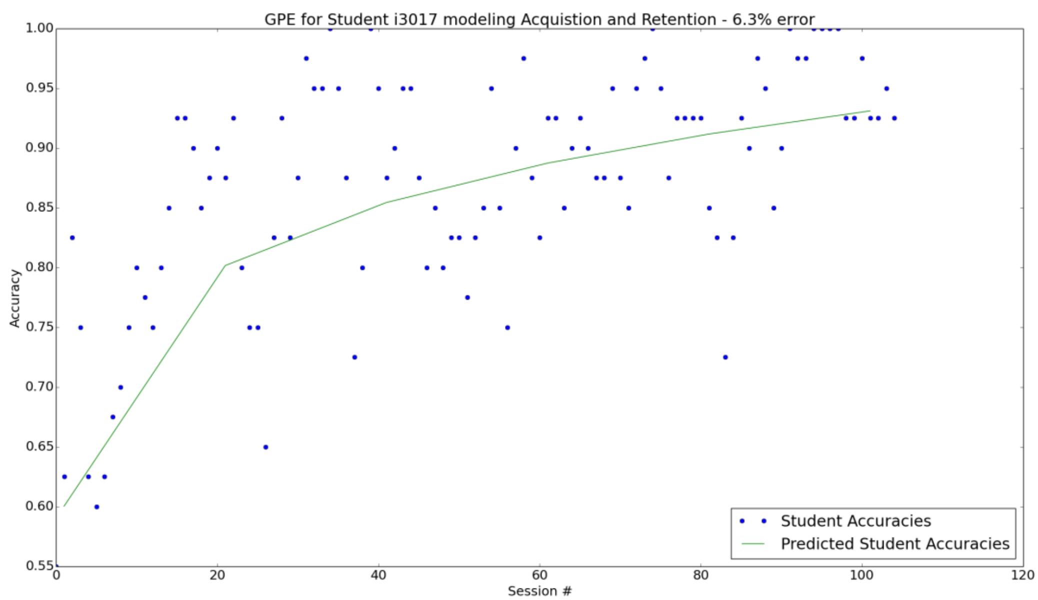

Figure 4: GPE (curve) and actual student performance (points) for student i3017 at difficulty level 4. The model shows continued but diminishing acquisition of skill from session to session.

Overall, 10 randomly selected students from the dataset were modeled and they produced an average MAD of 8.03% with a standard deviation of 1.06%.

Hidden Markov Models (HMM)

The HMM described in the models section is a sequential model that adapts its transition and observation probabilities as it is continually supplied with new performance data; thereby making it an advantageous model over the GPE. In this case, the sequence of accuracies that are observed from the student are used to infer a state (Novice, Intermediate, Expert) that the user may be in. Given a state, a Gaussian model is found that exhibits the observations at the stage of the sequence. The model outputs the most likely sequence of observations exhibited by the user, given the user's past observations. The input to this model is the sequence of observations for the same set of data that was used to build the GPE model. Figure 5 shows the performance of the HMM predicting the sequence of accuracies as compared to the actual sequence.

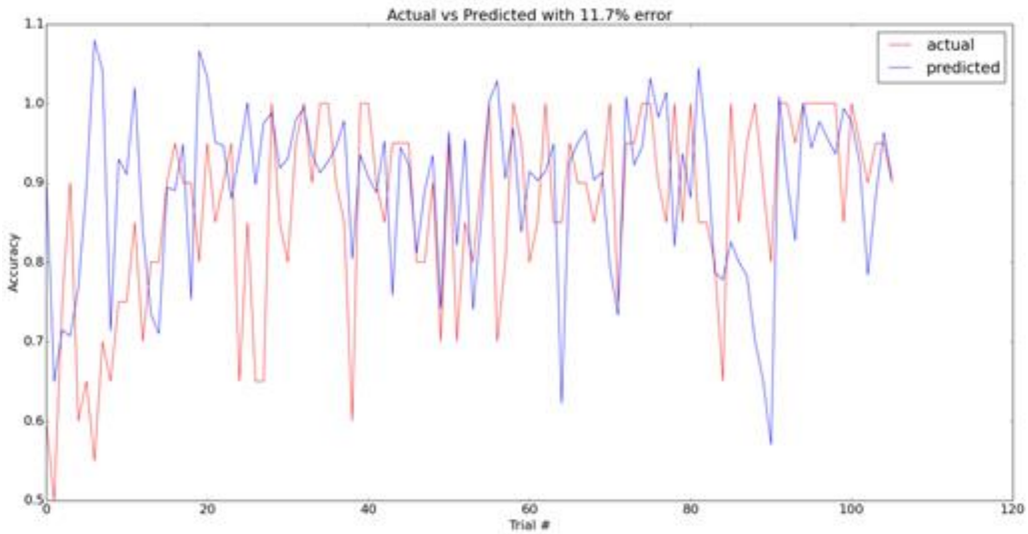

Figure 5: HMM predictions of a sequence of accuracy values over trials (in blue) err by 11.7% from actual accuracies (in red).

It should be noted that the HMM takes in the sequence of observations to produce a sequential output that predicts the user's performance at each session. This allows the HMM to adapt its output sequence based on the current instance without assuming a fixed form for the model. However, the use of only sequential information adds one shortcoming to the model. The concept of time is lost in this model and the inputs are interpreted as an ordered sequence of trials at a specific difficulty level (see the x-axis). Therefore, this model is appropriate when training sessions occur at uniform time intervals; it is not generalizable to training sessions that have a high variance in delays between them. In addition, the HMM will not make very accurate predictions for sessions that are held at times beyond that on which the model was trained.

Multi-Layer Perceptron Models (MLP)

A generalized temporal model is needed to account for prior information, capitalize on the advantages of the HMM, and use time as a ratio measure (rather than as an ordering of events). The MLP described in the modeling section accounts for the needs beyond that of the GPE and HMM. As a true predictor model, the data from 10 of the participants was left out of the model training set and used for model validation. The test phase of this "train, test" method of model performance analysis and validation represents how the developed models would perform on new participants in a real-world use case and scenario. The model is generalized over 231 participants, and is a temporal model that is able to easily input all the available features within the dataset to account for all prior information to make its predictions. MLPs developed for this effort included a history of 20 trials and as such included a total of 8 inputs (20 past sessions x 4 features/session + 4 features for current session, where the features include: task difficulty, time delta (i.e. time between sessions), historical performance, trial duration). When historical data was not available as inputs (i.e. trainees had not completed 20 historical sessions), model inputs were defaulted to "0" values. Iterative MLPs were validated using prediction horizons (i.e. number of future sessions to generate learning curve estimate) of 1, 3, 5, 10, and 15 sessions. The fixed trajectory MLPs developed for this effort implemented prediction horizons of 5 and 10.

The performance of the iterative MLP was evaluated across five prediction horizons. As expected, the iterative MLP was most accurate using a prediction horizon of 1 future session. However, error did not significantly increase for prediction horizons > 1. Error of model-generated predictions (expressed as MAD%) was calculated as 24.2%, 27.8%, 28.7%, 28.6%, and 28.4% for prediction horizons of 1, 3, 5, 10, and 15 sessions respectively (see Figure 6 for PH 1 form a specific student). An interesting finding of this validation was that it did not appear that an increased prediction horizon had a significant effect on overall model accuracy. One of the potential limitations of this modeling approach is that if model outputs (used iteratively as inputs) are erroneous, future predictions have a corresponding increase in error after each, subsequent iteration. Participants 111, 196, and 213 had the worst model performance. Learning by these participants differed significantly from that of the larger group that the model was trained to estimate. Participant 115 had a limited dataset (<10 trials), and as such, predictions could not be generated for models with prediction horizons >=10. Over all 10 participants, error (expressed as MAD%) was calculated as 28.4% for the largest prediction horizon.

The fixed trajectory MLPs developed in this effort were found to have better predictive accuracy than the iterative MLP models summarized above. We hypothesize that this is likely attributed to the fact that MLP outputs are fixed and do not have the potential sources of error that model inputs used in the iterative MLP (i.e. where model outputs/predictions are used as inputs as described above). Overall, MAD for the fixed trajectory MLP models developed for this effort was calculated as 17.3% and 18.7% for prediction horizons of 5 and 10 respectively when evaluated using data from the same 10 participants used for iterative MLP validation described above.

Figure 6 illustrates that the developed iterative MLP tracks trends in performance reasonably well for each participant. The developed MLP also appears to be sensitive to regions of "learning decay" which appear to have occurred sporadically across the participant population used for model training and validation. We hypothesize that the inclusion of time between sessions (in days) as a model input feature will help the model better predict these instances of decay. It appears there exist some instances where the model-generated estimates "lead" or are "delayed" with respect to the actual performance measures. Although these instances occur, a priori knowledge of significant changes in trainee learning curve that evolve over time may be sufficient to optimize curriculum.

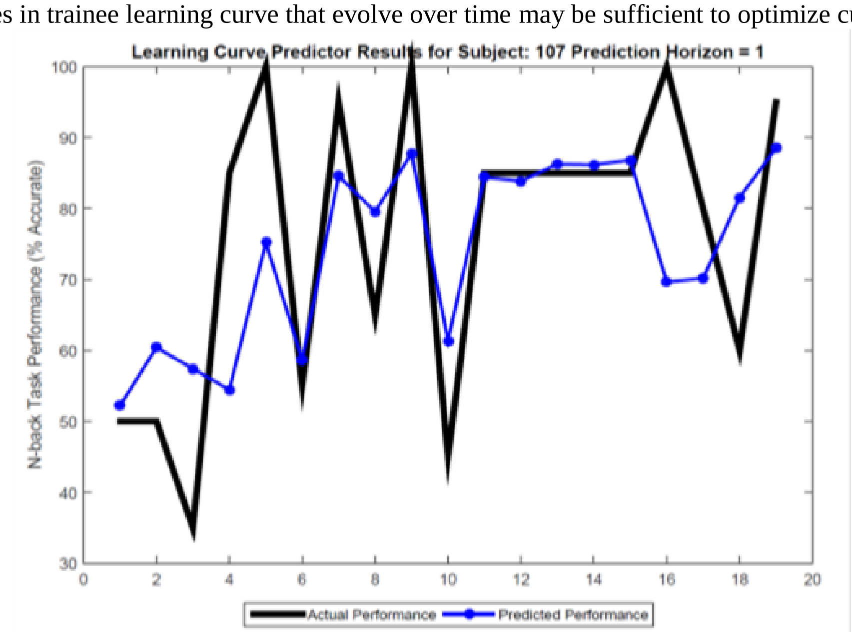

Figure 6: Iterative MLP model performance for participant 107 where the next point is predicted (in blue) based on all previous trials, and is shown to be comparable to the actual performance of the participant (in black).

CONCLUSION

The research reported in this paper explored methods of analyzing skill acquisition and decay using the traditional General Performance Equation, and two machine learning models: Hidden Markov Models and Multi-Layer Perceptron Models. The HMM very closely approached the accuracy of the GPE (see Table 1). The MLP-based models exhibited higher error in model predictions, but these models were evaluated across more extended prediction horizons (i.e. predicting performance >1 session ahead). Additionally, the iterative MLP-based models provide the additional capability to generate predictions over any desired prediction horizon.

Table 1: MAD error for the three models tested

	GPE	HMM	Iterative MLP (PH=1,3,5,10,15)	Fixed Trajectory MLP (PH = 5,10)
Error (average over 10 participants)	~8.03%	10.3%	27.5%	18.0%

We speculate that prediction accuracy may be significantly improved using additional data concerning training tasks, training objectives, and trainees themselves. This is one strength of the HMM and MLP. Specifically, they can easily incorporate features or variables that may have an effect learning acquisition and decay, as well as provide a method to identify said features or variables. For example, if we have a student's past scores on a supposedly unrelated task, the MLP provides a method to test whether these scores (and, thus, the tasks) are correlated with their current performance as well as incorporate them in the model to provide for more accurate long term predictions. Table 2 summarizes the advantages and disadvantages of all of the models tested here.

Table 2: Description, advantages, and disadvantages of the various modeling techniques investigated

Model	Short Description	Advantages	Disadvantages
GPE	Regression model assuming exponential learning and decay models	Well-researched and most accurate method for simple training tasks	Assumes fixed functional form and fits data to that form for all time and student performance.
HMM	Machine learning model assuming a student is in a state and models the transition between states as learning/decay with a set of observations (grades) from each state	Modeling explicit transitions between states of a student allows for incorporating different observation models in each state.	Temporal component implicitly assumes a fixed time delta between each transition. Prior information only includes single past state.
MLP	Machine learning model based on neural networks	Easily incorporates multiple features and other pertinent information that could affect student performance	Requires a lot of data for comparable performance to GPE

Future research in this area should examine the impact on forecasting accuracy of representing attributes of the student, domain, and learning environment. Future training technology development efforts should extend and integrate the capability, explored here, to estimate learning curves from performance data captured in training systems and task management systems. The goal is to use those estimates to recommend specific training or job experiences at specific times to accelerate acquisition and manage decay of key knowledge and skills. The effect will be a workforce whose skills are faster learned, better held, and better known to America's Armed Forces.

ACKNOWLEDGEMENTS

The authors thank Dr. Ray Perez of the Office of Naval Research for his support of the research reported here. The opinions expressed here are those of the authors, and not necessarily those of ONR, the Navy, or the Department of Defense.

REFERENCES

Anderson, J. R., and Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In Glaser, R. (ed.), *Advances In Instructional Psychology: Educational Design And Cognitive Science*. Mahwah, NJ: Erlbaum, pgs. 1-33.

Arthur Jr., Winfred et al. (1998). Factors That Influence Skill Decay And Retention: A Quantitative Review And Analysis. *Human Performance* 11.1: 57-101.

Ebbinghaus, H. (1999). *Memory. A Contribution To Experimental Psychology.*. New York: Thoemmes Press Reprint 1999, 1913.

Farr, Marshall J. (1987). *The Long-Term Retention Of Knowledge And Skills*. New York: Springer-Verlag.

Fitts, P. M., Posner, M.I. (1967). *Human Performance.* Oxford, England: Brooks And Cole.

Jastrzembski, T. S., K. A. Gluck, and G. Gunzelmann. (2006). Knowledge Tracing And Prediction Of Future Trainee Performance. In *Proceedings Of The Interservice/Industry Training, Simulation, And Education Conference (I/ITSEC)*. Orlando, FL: National Training Systems Association.

Lesgold, A. M. (2012). Practical Issues In The Deployment Of New Training Technology. In P. J. Durlach and A. M. Lesgold (eds.), *Adaptive Technologies For Training And Education*. New York: Cambridge University Press.

Pavlik, Philip I., and John R. Anderson. (2005). Practice And Forgetting Effects On Vocabulary Memory: An Activation-Based Model Of The Spacing Effect. *Cognitive Science* 29.4: 559-586.

Ward, N.; Paul, E.; Watson, P.; Cooke, G. E.; Hillman, C. H.; Cohen, N. J.; Kramer, A. F.; and Barbey, A. K. (In review). Illuminating the Limited Efficacy of Brain Training.

Wang, X., Day, E.A., Kowollik, V., Schuelke, M.J., and Hugues, M.J. (2013). Factors Influencing Knowledge And Skill Decay After Training: A Meta-Analysis. In W. Arthur Jr., et al. (eds.), *Individual and Team Skill Decay: The Science and Implications for Practice*. NY: Taylor & Francis Group.