A Dynamical System Model of Microgenetic Changes in Performance, Efficacy, Strategy Use and Value during Vocabulary Learning

P. PAVLIK JR. AND S. WU
Carnegie Mellon University, United States

_____

This paper describes the development of a dynamical systems model of self-regulated learning, which explains the practically and theoretically important dynamic relationships among three student-engagement constructs and performance during learning. This work mined data from computerized adaptive flashcard learning to create this dynamical systems model. This flashcard practice included pop-up survey questions on the student's experience of recent easiness, strategy use, and usefulness, in addition to the correctness performance data for the practice. Using this dynamical systems model, we were then able to simulate various user profiles to predict how they would experience the flashcard system. These simulations show how strategy use in this task is crucial because of the ways it influences performance, particularly over time. In the model, this result is shown by a bifurcation into two different regions for higher and lower strategy use, where the higher strategy-use equilibrium state is accompanied by performance predictions suggesting learning that is more efficient.

Key Words and Phrases: Dynamical system model, language learning, motivation, metacognition, efficacy, and utility

_____

1. INTRODUCTION

Many sources agree that motivation is a dynamic construct that changes from moment to moment with the stream of events students experience [e.g. Witherspoon et al. 2007]. Of course, while each period marks changes in student attitudes, perceptions and actions, there is also continuity across time as attitude states shift not randomly, but as a function of prior attitudes, perceptions and actions. If these different variables cause change in each other, in addition to this continuity, we have described a dynamical system. This perspective on dynamic motivation is similar to Csikszentmihalyi's definition of the state of flow that occurs for people during activities that balance skills and challenges [Csikszentmihalyi 1991]. This paper attempts to add richness to the discussion, by introducing a flow-like mathematical model that supposes there is more than just skill and challenge feeding into the changing experience of motivation in a student.

This work resonates well with the general project of dynamical systems research in the social and cognitive sciences [Vallacher and Nowak 2007; Ward 2002], by providing a tangible example of a dynamical system. However, unlike prior work, it appears to be more focused on a generalizable model with educational implications. In the one instance we were able to find of an educationally relevant dynamical model that was fit to data [Guastello et al. 1999], the authors chose to fit the system to each user, a procedure that produces a distribution of dynamical systems. While this procedure produces interesting results, it is more difficult to use for making general predictions.

In this paper, we data mine for our general dynamical system in a dataset from a Chinese vocabulary tutoring system. This task is superficially straightforward, because students see an "optimal" schedule of drill practice (test with a review if incorrect) and have to respond to each drill practice by typing the matching Chinese pinyin phonetic representation or English meaning. Additionally, because we were pursuing the hypothesis that strategies and motivation might influence this practice, we asked three

_____

Authors' addresses: P. Pavlik Jr, Human Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Email: ppavlik@andrew.cmu.edu; S. Wu, Carnegie Mellon University, Pittsburgh, PA, USA. Email: suemei@andrew.cmu.edu

questions, which cycled for each user (using individual random orders for each subject, i.e. Q1, Q2, Q3 or Q2, Q1, Q3, etc.…). One of these questions popped up every 2 minutes during the flashcards, and required only a single key response to minimize student inconvenience. These questions were as follows:

- "How easy was the recent practice?", with Likert like items ranging from "too hard" to "too easy", on a 5 point scale.
- "How useful for learning was the recent practice?", with Likert like items ranging from "not useful" to "very useful", on a 5-point scale.
- "Were you able to use any learning strategies during the recent practice?" with Likert like items ranging from "mostly used repetition" to "mostly used strategies", on a 5 point scale.

These questions were meant to be unthreatening to the student, since they are only admissions about recent practice, and so should not carry strong social norms for responding one way or another. These questions attempt to measure 3 crucial theoretical determinants of performance and motivation that have been suggested over the last 50 years in the literature on behavior, motivation, and metacognition.

2. DATA COLLECTION
The data was collected in the Elementary Chinese II course at Carnegie Mellon University (CMU). All data used to create the model have been archived in the Pittsburgh Science of Learning Center (PSLC) DataShop web application, and a copy of the model and data are available to registered users (free). The R code posted at this location functions by downloading the data using DataShop web services (requires a quick set-up) before finding the model using R functions as described below.

The data itself comes from Lessons 15-18 in Elementary Chinese II at Carnegie Mellon University, which correspond to chapters 15-18 in the *Chinese Link* (*Zhongwen Tiandi*) textbook [Wu et al. 2006]. The intervention was a computerized adaptive practice drill system with several different types of Chinese vocabulary flashcards including practice of the pinyin phonetic system, English meaning, radical character components of Chinese, and flashcards where the student filled in a missing vocabulary word in a sentence (Cloze fill in the blank) [Pavlik Jr. et al. 2008]. We had 65 unique subjects (of which only 61 produced complete practice data) and 198 unique lesson runs. Average lesson-run length was about 20 minutes (averaging 9.34 blocks of 2 minutes of practice per lesson started). Students were allowed to take breaks (by hitting a pause button) at any time, the dynamical systems model ignores any breaks students took and treats each lesson as continuous. Since students were asked to practice for only 20 minutes, many students completed the work without any break. While there was a between-subject manipulation of initial practice difficulty (either wider or narrower distribution of initial practice), this did not appear to cause detectable effects, though it may have produced more varying data, thus providing more variability for the model to capture, improving the detection of patterns by encouraging the occurrence of these patterns in the data.

3. DYNAMICAL SYSTEM MODEL
After doing some preliminary correlations of the t and t-1 data values (using only one predictor and one predicted variable) to make sure there were significant relationships between the variables, the data was organized into a set of data points for each user for each lesson. There were 4 data vectors per student per lesson, including vectors for easiness, usefulness, strategy use, and performance. Since each of the 3 survey items was measured sparsely, about 2/3 of the data values in each survey vector were left empty. Probability correct for the previous 2 minutes (previous epoch) however, was always

calculable, so these vectors were not sparse. Further, since probability (performance) is a 0 to 1 value (tending to be greater than 50% in the data), whereas our readings from the other measures varied from 1-5, we scaled probability by using the logit of probability as data (constrained so that logit<-5 =-5 and logit>5 =5). For this reason, the model predicts logit values for performance, and the simulator code converts these to probabilities for the simulation graph plotting. Because the model is undefined unless it has a full set of 4 prior values to begin computations, we used the neutral values of 3 easiness, 3 usefulness and 3 strategy use to initialize the model for each lesson for each user.

A dynamical systems model assumes some current state of nature and then describes an evolution rule about how each state is transformed to the future state. Assuming the evolution rule is correct; by iteration, we can predict the future states of nature given some start state. Our evolution rule was

$$e_{t+1} = x_1 e_t + x_2 e_t^2 + x_3 u_t + x_4 u_t^2 + x_5 s_t + x_6 s_t^2 + x_7 p_t + x_8 p_t^2 +$$
$$x_9 e_t u_t + x_{10} e_t s_t + x_{11} e_t p_t + x_{12} u_t s_t + x_{13} u_t p_t + x_{14} s_t p_t + x_{15}$$

$$u_{t+1} = x_{16} e_t + x_{17} e_t^2 + x_{18} u_t + x_{19} u_t^2 + x_{20} s_t + x_{21} s_t^2 + x_{22} p_t + x_{23} p_t^2 +$$
$$x_{24} e_t u_t + x_{25} e_t s_t + x_{26} e_t p_t + x_{27} u_t s_t + x_{28} u_t p_t + x_{29} s_t p_t + x_{30}$$

$$s_{t+1} = x_{31} e_t + x_{32} e_t^2 + x_{33} u_t + x_{34} u_t^2 + x_{35} s_t + x_{36} s_t^2 + x_{37} p_t + x_{38} p_t^2 +$$
$$x_{39} e_t u_t + x_{40} e_t s_t + x_{41} e_t p_t + x_{42} u_t s_t + x_{43} u_t p_t + x_{44} s_t p_t + x_{45}$$

$$p_{t+1} = x_{46} e_t + x_{47} e_t^2 + x_{48} u_t + x_{49} u_t^2 + x_{50} s_t + x_{51} s_t^2 + x_{52} p_t + x_{53} p_t^2 +$$
$$x_{54} e_t u_t + x_{55} e_t s_t + x_{56} e_t p_t + x_{57} u_t s_t + x_{58} u_t p_t + x_{59} s_t p_t + x_{60}$$

where t – Indexes the observation time epoch (t+1 values are predicted values)
e – Likert-like self-report on recent easiness
u – Likert-like self-report on recent usefulness
s – Likert-like self-report on recent strategy use
p – logit of probability correct since previous self-report (max 5 min -5)

This rule was applied for each of the 4 state space variables to compute the next epoch prediction. The rule for each variable required 15 parameters to capture the 4 first order (linear) effects, 10 second order (quadratic) effects, and the single fixed (intercept) effect. This meant we had 60 parameters for the non-linear deterministic model.
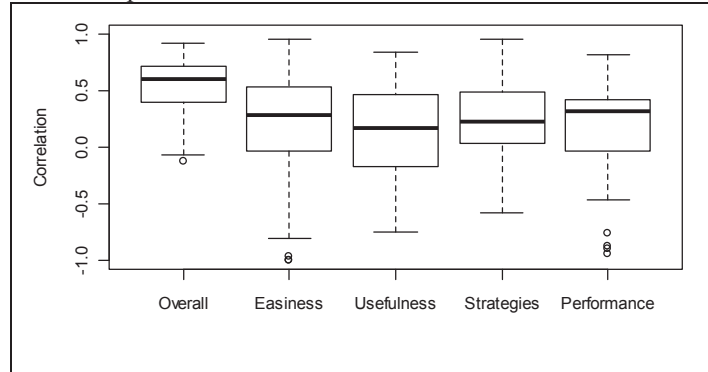

Fig. 1. Boxplots of correlations between model and data for each individual student, for the overall model and by construct (N=61 students).

The model parameters were found by fitting the model to the data by minimizing the sum of squared error. To fit the model we used the BFGS optimizer implementation in the R "optim" command. Because there was not sufficient time to complete some sort of model validation, these results should be considered preliminary. However, to provide some support for validity, Figure 1 shows boxplots of the correlations of the model and data for each individual student, for each of the 4 variables. These correlations show the model fit is not achieved by fitting individual student magnitudes, but rather the model is capturing a general overall model of the dynamical system within the bulk of students. In

particular, it is encouraging that only 2 students negatively correlated with the overall model. It is also clear the model is not succeeding by just fitting one or some of the constructs while neglecting others. In fact, we see the overall pattern is better modeled than any individual construct.

5. SIMULATIONS

We wanted to use the model as a simulation tool to show us how to understand the student experiences in the tutor and how these experiences are reflected in the constructs we tracked.
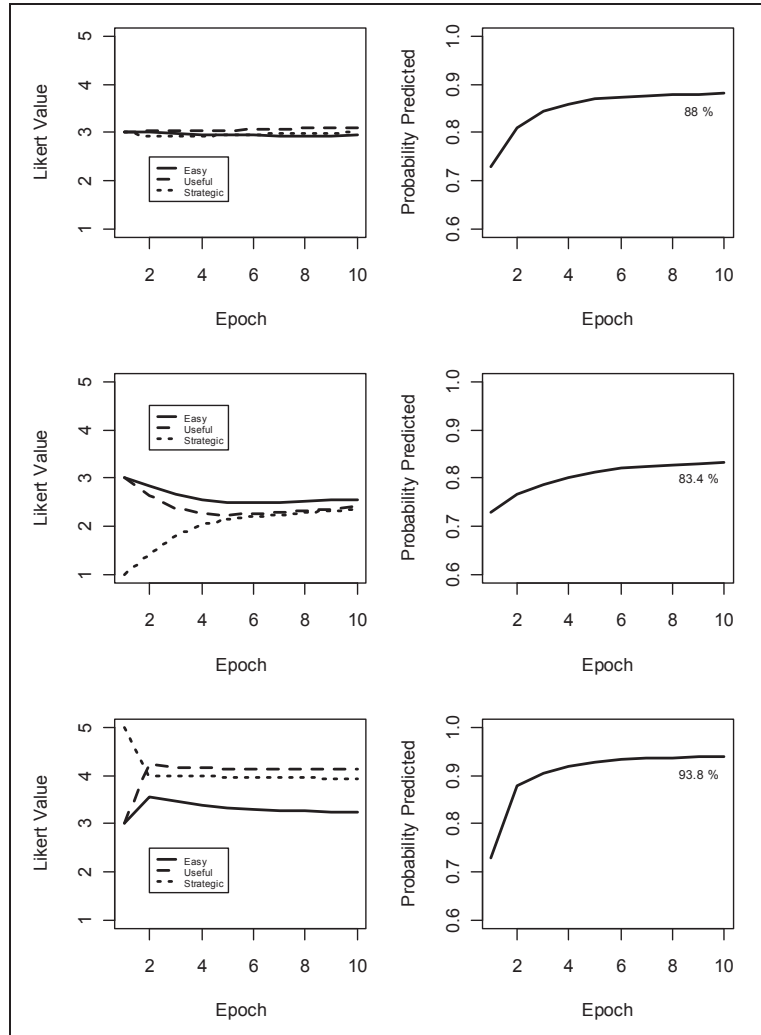

Fig. 2. Pinned simulations of an average student, a rote learner, and a strategic learner.

To begin, Figure 2 top shows a pinned (initialized at a certain value) simulation starting from values very close to average (we used 3,3,3 for simplicity and because these were clearly the medians of each distribution). For this example, the model was initialized with 3 easiness, 3 usefulness, 3 strategy-use, and a starting performance of 73% (equivalent to a logit of 1, which was almost exactly the true average). Figure 2 top demonstrates this

"average" model reaches a performance level of 88.0% after 20 minutes (10 epochs) of practice, and suggests visually this level of performance has reached a stable plateau.

Figure 2 middle for a rote learner contrasts with Figure 2 bottom, for a strategic learner, each pinned to begin at the minimum and maximum strategy values. This strategic simulated student has a much different experience and we can see after 10 epochs, it is already becoming clear they have arrived at a different steady state with strategy use and usefulness ratings stabilizing at about 4. Interestingly, easiness is not improved, in large part because (according to our parameter analysis, not shown) high strategy use and high usefulness combined are predicted to reduce easiness. This makes sense if we suppose producing strategies in a useful fashion might be difficult for the student, even if effective. This effectiveness is highlighted by the reduced error rate in this steady state high strategy model. Comparing the error rates (1-.938=.062) for pinned high and pinned low (1-.834=.166) reveals pinning to the low strategy for the first epoch leads to the prediction of .166/.062=2.7 times more errors during the epoch 16 minutes later. Because errors greatly reduce practice efficiency in flashcard type learning [Pavlik Jr. and Anderson 2008], such an effect shows how strategy use may be an important component of optimal practice.

6. CONCLUSIONS

The main contribution of this work is to provide a general model to help us understand the dynamic motivational system of students. The dynamical systems model we produced is valuable because of what it implies for optimal use of the flashcard system and optimal learning in general. As we showed in the simulations, the model provides evidence of the mutually supportive feedback between usefulness, strategy use, and performance. By observing the simulations we see how high strategy use has a powerful effect on students' perception of usefulness, but this is contrasted by the different behaviour of easiness ratings, which are depressed when strategies and usefulness are high. This seems to be a natural result of people finding it less easy to use strategies despite the acknowledgment of their usefulness for learning.

ACKNOWLEDGEMENTS

REFERENCES

CSIKSZENTMIHALYI, M. 1991. *Flow: The Psychology of Optimal Experience*. Harper Perennial.

GUASTELLO, S.J., JOHNSON, E.A. AND RIEKE, M.L. 1999. Nonlinear Dynamics of Motivational Flow. *Nonlinear Dynamics, Psychology, and Life Sciences 3*, 259-273.

PAVLIK JR., P.I. AND ANDERSON, J.R. 2008. Using a model to compute the optimal schedule of practice. *Journal of Experimental Psychology: Applied 14*, 101-117.

PAVLIK JR., P.I., PRESSON, N. AND HORA, D. 2008. Using the FaCT System (Fact and Concept Training System) for Classroom and Laboratory Experiments. In *Inter-Science Of Learning Center Conference*, Pittsburgh, PA.

VALLACHER, R.R. AND NOWAK, A. 2007. Dynamical social psychology: Finding order in the flow of human experience. In *Social Psychology: Handbook of Basic PrinciplesConference Name*, A.W. KRUGLANSKI AND E.T. HIGGINS Eds. Guildford Publications, New York.

WARD, L.M. 2002. *Dynamical cognitive science*. MIT Press, Cambridge, Mass.

WITHERSPOON, A., AZEVEDO, R., GREENE, J., MOOS, D. AND BAKER, S. 2007. The Dynamic Nature of Self-Regulatory Behavior in Self-Regulated Learning and Externally-Regulated Learning Episodes. In *Proceedings of the 13th InternationalConference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That WorkConference Name*, R. LUCKIN AND K.R. KOEDINGER Eds. IOS Press, 179-186.

WU, S.-M., YU, Y. AND ZHANG, Y. 2006. *Chinese Link: Zhongwen Tiandi, Intermediate Chinese*. Pearson Education/ Prentice Hall.