Team Neurodynamics
by The
Learning Chameleon, Inc.

Integrating EEG Models of Cognitive Load with Machine Learning Models of Scientific Problem Solving.

Ron. Stevens

IMMEX Project, UCLA School of Medicine
5601 W. Slauson Ave. #255, Culver City, CA 90230
immex_ron@hotmail.com

Trysha Galloway

IMMEX Project, UCLA School of Medicine
5601 W. Slauson Ave. #255, Culver City, CA 90230
TryshaG@gmail.com

Chris Berka

Advanced Brain Monitoring
2850 Pio Pico Dr. Ste. A,
Carlsbad, CA 92008
chris@b-alert.com

Abstract

We have conducted studies with wireless headsets to explore the relationship of working cognitive load (EEG-WL), distraction (EEG-DT) and engagement (EEG-E) and problem solving efficiency and effectiveness on a series of qualitative chemistry, biology and mathematics simulations. As students gained experience by working multiple cases, the EEG-E levels decreased with the reduced novelty of the problem space, the WL levels remained similar and the DT levels were variable. This EEG-DT variability was associated with the relative difficulty of the problem, by the misinterpretation of data, and / or by uncertainty associated with the solution of the problem.

To refine the analysis, real-time estimates of EEG-WL, EEG-DT and EEG-E obtained at one second intervals were interleaved with timeline representations of the problem solving process to associate the dynamics of cognitive function with the dynamics of problem solving and learning. Elevated EEG-E frequently occurred shortly after the selection of test data, especially during the first problem performance, and also when the students were taking notes. EEG-WL values fluctuated over the case performance but with no obvious relationship to EEG-DT or EEG-E. EEG-DT was closely linked with missing the case solution and was often reciprocal to the EEG-E. These results indicate that real time monitoring of EEG can begin to contribute a dynamic dimension to classroom problem solving and could help design approaches for real-time feedback to improve learning.

1 INTRODUCTION

Skill development has been described as occurring in stages that are characterized by distinctive amounts of time and mental effort required to exercise the skill (Anderson, 1982, 1995, Schneider and Shiffrin, 1977). Three stages have been identified: the initial cognitive stage requiring the assembling of new knowledge, the associative stage where newly assembled procedural steps gradually automate as they are practiced, and the autonomous stage where the task execution is automated and performed with minimal conscious mental effort. During the transition from the cognitive to associative stage, both speed and accuracy increase as subjects become less reliant on the declarative representations of knowledge (Anderson, 1982, 1995).

Given the complexities of skill acquisition it is not surprising that a variety of approaches have been used to develop models of the process. For instance, some researchers have explored the improved powers of computation in combination with machine learning tools to provide refined models of student skill acquisition and learning behaviours in science and mathematics. The scope and depth of the learning activities probed by these technology-driven tools are becoming increasingly detailed (Arroyo, Woolf & Beal, 2006; Beal, 2004). Such systems rely on learner models that include continually updated estimates of students' knowledge and misconceptions, based on actions such as choosing an incorrect answer or requesting a multimedia hint. Although such learner models are capable of forecasting student difficulties (Stevens, Johnson, & Soller, 2005), they still rely on relatively impoverished input.

Because the key distinction between the second and third stages of skill acquisition is a decrease in mental effort rather than a reliable difference in the accuracy of performance, application of neurophysiologic approaches, including the quantification of EEG correlates of workload, attention and task engagement have also been used to provide objective evidence of the progression from stage 2 to stage 3 (Berka, 2004, Berka 2006). If performance alone is used, no differentiation is made between people who perform well but require high mental effort and people who perform well with low mental effort.

Although there is a large and growing literature on the EEG correlates of attention, memory, and perception (Fabiani, 2001) there is a relative dearth of EEG investigations of the process skill acquisition and learning (Smith,1999). EEG researchers have generally elected to employ study protocols that utilize training-to-criterion to minimize variability across subjects and ensure that stable EEG parameters could be characterized. In most studies, the EEG data is not even acquired during the training process leaving a potentially rich data source untapped.

However, while advanced EEG monitoring devices are becoming more common in high workload / high stress professions (such as tactical command, air traffic control) the ideas have not been comprehensively applied to real-world educational settings due in part to some obvious challenges. First, the acquisition of problem solving skills is a gradual process and not all novices solve problems in the same way, nor do they follow the same path at the same pace as they develop an understanding of the domain. Next, given the diversity of the student population it is difficult to assess what their relative levels of competence are when performing a task making it difficult to accurately relate EEG measures to other measures of task skill. This is further complicated as strategic variability makes analyzing the patterns of students' problem solving record too complicated to be performed routinely.

Nevertheless, there are aspects of science education that could benefit from deriving data from advanced monitoring devices and combining it with real-time computational models of the tasks and associated outcomes. For instance, they could help identify where learning gains are being impeded by the task complexity or the need for the integration of diverse skills. Also, probabilistic models of student performance, and predictive extrapolations from these models to future performance are just that, probabilistic, lacking in the specifics of the individual. We believe that by combining such probabilistic models with the highly individualized cognitive measures afforded by EEG, the predictive accuracy of both models can be greatly improved and suggest approaches for augmenting cognition in educational settings. This paper describes our initial steps towards this goal.

2 METHODS

2.1 The IMMEX™ Problem Solving Environment

The software system used for these studies is termed IMMEX™ whose program structure is based on an extensive literature of how students select and use strategies during scientific problem solving (VanLehn, 1996, Schuun & Reder, 2001, Schuun et al., 2001, Haider & Frensch, 1996).

To illustrate the system, a sample biology task called *Phyto Phiasco* provides evidence of a student's ability to identify why the local potato plants are dying. The problem begins with a multimedia presentation explaining the scenario and the student's challenge is to identify the cause. The problem space contains 38 menu items describing local weather conditions, soil nutrients, plant appearance, disease symptoms, etc. When the student selects a menu item, he or she verifies the test requested and is then shown a presentation of the test results. When students feel they have gathered the information needed to identify the cause they attempt to solve the problem.

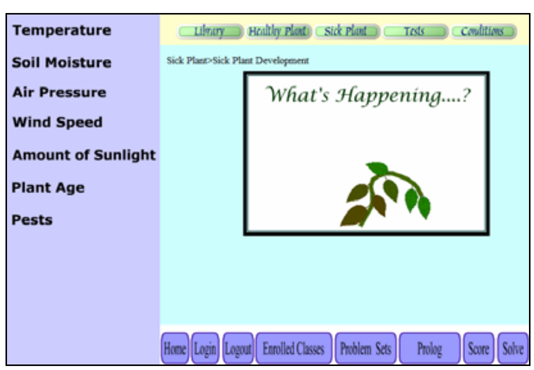

Figure 1. Sample IMMEX™ simulation. In the *Phyto Phiasco* simulation, the farmer's potato plants are dying and the challenge for the student is to identify the cause by examining local weather conditions, nutrients, etc.

The IMMEX™ database serializes timestamps of how students use these resources. Then in real time models are formed based on 1) estimates of student ability, 2) the strategies used, and 3) estimates of future performance. These are performed by Item Response Theory (IRT) analysis, Artificial Neural Network (ANN) analysis and Hidden Markov Modeling (HMM) respectively (Stevens et al., 2004, 2005, 2006).

For IRT analysis the problem difficulty of the different cases is first estimated using the solve rates from a large number of student performances. Then, using this model, the ability of each student is estimated using not only whether or not the case was solved, but also the relative difficulty of the case.

As students solve IMMEX™ cases, the menu items selected are then used to train competitive, self-organizing ANN (Stevens & Najafi, 1993, Stevens et al, 1996). Self-organizing maps learn to cluster similar performances in such a way that neurons near each other in the neuron layer respond to similar input vectors (Kohonen, 2001). The result is a topological arrangement of performance clusters where geometric distance between these clusters becomes a metaphor for strategic similarity. We frequently use a 36-node neural network and train with between 2000-5000 performances derived from students with different ability levels and where each student performed at least 3-4 cases of the problem set (Stevens & Casillas, 2006). The components of each strategy in this classification can then be visualized for each of the 36 nodes by histograms showing the frequency of items selected (Figure 2).

Most strategies defined in this way consist of items that are always selected for performances at that node (i.e. those with a frequency of 1) as well as items that are ordered more variably. For instance, many Node 15 performances shown in Figure 2 A contain the items 28-31 whereas few contain items 8-12. Figure 2 B is a composite ANN nodal map, which illustrates the topology generated during the self-organizing training process. Each of the 36 graphs in the matrix represents one node in the ANN, where each individual node summarizes groups of similar student's problem solving performances automatically clustered together by the ANN procedure.

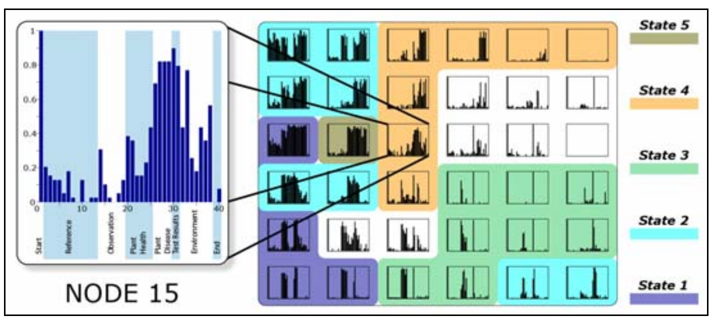

Figure. 2. Sample neural network nodal analysis for identifying strategies. a.) The selection frequency of each action (identified by the labels) is plotted for the performances at node 15, thus characterizing the performances for this node and relating them to performances at neighboring nodes. The nodes are numbered in rows, 1-6, 7-12, etc. b.) This figure shows the item selection frequencies for all 36 nodes (Stevens et al., 2004, 2005, 2006).

IMMEX™ problem sets consist of 5-60 parallel cases, so that if students perform multiple cases of a problem set, learning trajectories can be developed through Hidden Markov Modeling (HMM) that not only reflect and model students' strategy shifts as they attempt series of cases, but also predict future problem solving performance. In our context, a number of hidden states are postulated to exist that represent the strategic transitions that students may pass through as they perform multiple IMMEX™ cases. Each hidden state is represented by a probabilistic machine learning function, realizing the idea that student problem solving activities contain an element of uncertainty. For most problem sets, the postulated number of states is 3-5 based on the observed strategic complexity. Then, similar to ANN analysis, exemplars of strategy sequences, as identified by ANN are repeatedly presented to the HMM software to develop progress models. As shown in Figure 3, the emission matrix resulting from the HMM procedure allows a mapping to the different strategies encompassed by each state.

The dynamics of the state changes for *Phyto Phiasco* learning trajectory are shown in Figure 3. Initially many students began by selecting many test items as represented by State 2 and consistent with models of skill acquisition (Ericsson, 2004), with time they refined their strategies and selected fewer tests as shown by States 3 and 4. As expected, with practice student's solve rates increased from 35% to 63% ($\chi2 = 121.8$, df=10, $p<0.000$). The rate of stabilization, and the strategies stabilized are influenced by gender (Stevens & Soller, 2005), experience (Stevens et al, 2004), and individual or group collaboration (Cooper et al, submitted), etc. Students often continue to use these stabilized strategies for prolonged periods of time (3-4 months) when serially re-tested (Stevens, 2006).

IMMEX™ problem solving therefore represents a task where it is possible to construct probabilistic models of many different aspects of problem solving skill acquisition across problem solving domains. The constraints of working memory are likely to be particularly relevant during such skill acquisition where working memory capacity can frequently be exceeded. The possibility of combining these models with EEG workload metrics raises questions regarding student learning: e.g., what are the relative cognitive demands and the balances of different working memory capacities as students gain experience and begin to stabilize their strategies?

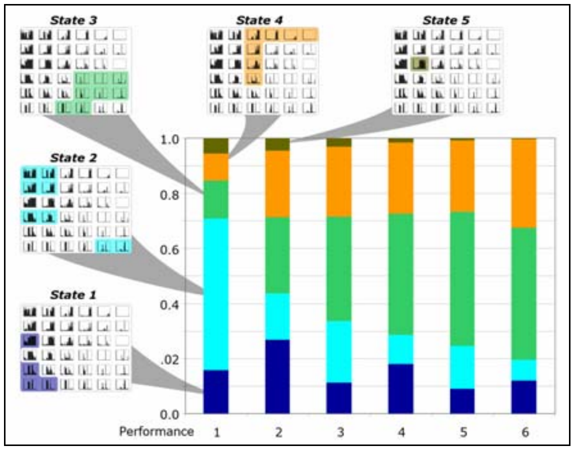

Figure 3. Modeling individual learning trajectories. This figure shows the strategic changes as students working alone gain experience in *Phyto Phiasco* problem solving. Each stacked bar shows the distribution of HMM states for the students (N=3325) after a series (1-6) of performances. These states are also mapped back to the 6 x 6 matrices which represent 36 different strategy groups identified by self organizing ANN. The highlighted boxes in each neural network map indicate which strategies are most frequently associated with each State (Stevens et al. 2004).

2.2 The B-Alert®system

Recording and analysis of EEG has traditionally been confined to laboratory settings due to the technical obstacles of recording high quality data and the computational demands of real-time analysis. Advances in electronics and data processing set the stage for ambulatory EEG applications. A recently developed wireless EEG sensor headset facilitates easy acquisition of high quality EEG combining battery-powered hardware with a sensor placement system to provide a lightweight, easy-to-apply method to acquire and analyze six channels of high-quality EEG.

The EEG sensor headset requires no scalp preparation and provides a comfortable and secure sensor-scalp interface for 12 to 24 hours of continuous use. The headset was designed with fixed sensor locations for three sizes (e.g., small, medium and large). Standardized sensor placements include locations over frontal, central, parietal and occipital regions (sensor sites: F3-F4, C3-C4, Cz-PO, F3-Cz, Fz-C3, Fz-PO). Amplification, digitization, and radio frequency (RF) transmission of the signals are accomplished with miniaturized electronics in a portable unit worn on the head. The combination of amplification and digitization of the EEG close to the sensors and wireless transmission of the data facilitates the acquisition of high quality signals even in high electromagnetic interference environments. Data are sampled at 256 samples/second with a bandpass from 0.5 Hz and 65Hz (at 3dB attenuation) obtained digitally with Sigma-Delta A/D converters.

Quantification of the EEG in real-time, referred to as the B-Alert® system, is achieved using signal analysis techniques to identify and decontaminate fast and slow eye blinks, and identify and reject data points contaminated with excessive muscle activity, amplifier saturation, and/or excursions due to movement artifacts. Decontaminated EEG is then segmented into overlapping 256 data-point windows called overlays. An epoch consists of three consecutive overlays. Fast-Fourier transform is applied to each overlay of the decontaminated EEG signal multiplied by the Kaiser window ($\alpha = 6.0$) to compute the power spectral densities (PSD). The PSD values are adjusted to take into account zero values inserted for artifact contaminated data points.

Wavelet analyses are applied to detect excessive muscle activity (EMG) and to identify and decontaminate eye blinks. Once the artifacts are identified in the time-domain data, the EEG signal is decomposed using a wavelets transformation. The wavelets eye blink identification routine uses a two-step discriminant function analysis (DFA).

The DFA classifies each data point as a control, eye blink or theta activity. Multiple data points that are classified as eye blinks are then linked and the eye blink detection region is established. Decontamination of eye blinks is accomplished by computing mean wavelet coefficients for the 0-2, 2-4 and 4-8 Hz bins from nearby non-contaminated regions and replacing the contaminated data points. The EEG signal is then reconstructed from the wavelets bins ranging from 0.5 to 64 Hz. Zero values are inserted into the reconstructed EEG signal at zero crossing before and after spikes, excursions and saturations. EEG absolute and relative power spectral density (PSD) variables for each 1-second epoch using a 50% overlapping window are then computed.

2.3 Subjects and Study

Subjects (n=7) first performed a single 30-minute baseline EEG test session to adjust the software to accommodate individual differences in the EEG (Berka, 2004). They then performed multiple IMMEX™ problem sets targeted for 8th-10th grade students. These include *Phyto Phiasco*, the biology problem described above, *Get Organized* where the goal is to diagnose disorders of organ systems, and a mathematics problem called *Paul's Pepperoni Pizza Palace*. Subjects generally performed at least 3 cases of each problem set allowing the tracking of changes in EEG-DT, EEG-E and EEG-WL across cases, as well as across problem sets as students gain experience. In the second part of the study we aligned the EEG output metrics on a second-by-second basis with the problem solving actions to explore the within-task EEG metric changes. The total session time lasted about 2.5 hours.

The output of the B-Alert software includes EEG metrics (values ranging from 0.1-1.0) for engagement, distraction and workload calculated for each 1-second epoch of EEG using quadratic and linear discriminant function analyses of model-selected EEG variables derived from power spectral analysis of the 1-Hz bins from 1-40Hz. These metrics have proven utility in tracking both phasic and tonic changes in cognitive states, in predicting errors that result from either fatigue or overload and in identifying the transition from novice to expert during skill acquisition. (Berka, 2004, Berka 2005).

3 RESULTS

3.1 Dynamics of EEG-WL, EEG-DT and EEG-E with Practice

We first examined the changes in EEG-E, EEG-DT and EEG-WL as students performed their first and second cases of different problem sets (n=15) as this is where the greatest shifts in strategy normally occur (Figure 3) (Table 1). Only the EEG-E showed significant changes between the two performances decreasing ~20%.

Table 1. EEG Metrics Across Cases 1 and 2. The average EEG-E, EEG-DT and EEG-WL were compared for 7 sets of performances of the IMMEX simulations.

	Case 1	**Case 2**	**p**
EEG-E	43 ± 11	36 ± 9	.04
EEG-DT	13 ± 8.8	22 ± 6	.302
EEG-WL	56 ± 11	38 ± 11	.132

We then averaged the EEG-Dist, EEG-LE and HE and EEG-HWL during the interval between the test selections. These intervals ranged from 11 sec. to 50 sec. As shown in Table 2 the average Brain State varied over a wide range (.52-.69) across the intervals suggesting that a finer analysis, perhaps combined with a more detailed analysis of the actions of the student may begin to reveal more subtle aspects of the problem solving process. For example, peak WL was observed during the 2nd and 4th periods suggesting increased problem solving or analysis.

Table 2. Inter Test Averages of EEG – Measures.

	Interval	B. State	%SO	%Dist	%LE	%HE	%HWL	N =
all aver		0.60	0.01	0.26	0.44	0.29	0.63	291.00
1 aver	47	0.56	0.00	0.33	0.38	0.29	0.68	46.00
2 aver	19	0.49	0.06	0.35	0.31	0.29	0.59	19.00
3 aver	20	0.57	0.00	0.31	0.52	0.17	0.69	19.00
4 aver	11	0.61	0.04	0.19	0.47	0.30	0.61	12.00
5 aver	50	0.69	0.01	0.17	0.45	0.37	0.60	49.00
6 aver	14	0.64	0.00	0.31	0.42	0.27	0.58	14.00
7 aver	29	0.56	0.00	0.35	0.43	0.22	0.63	54.00
8 aver	26	0.52	0.03	0.35	0.37	0.25	0.60	18.00
9 aver	18	0.60	0.01	0.17	0.54	0.27	0.61	29.00
10 aver	37	0.65	0.00	0.16	0.47	0.37	0.64	12.00

The average maximum percentage values for Brain State (B. State), Sleep State (SO), Distraction (Dist), Low Engagement (LE), High Engagement (HE) and Work Load (WL) were calculated for the intervals between different test selections. Values above the average are highlighted in red.

While there were no differences in the EEG DT values across the first and second performances of the different problem sets, post-session interviews suggested that higher EEG-DT levels might be associated with the difficulty of the case. Support for this suggestion is shown in Figure 4 where the IRT item difficulties of the different cases being performed are plotted against the EEG-DT levels. The correlation ($R^2 = .63$) was significant ($p = 0.03$).

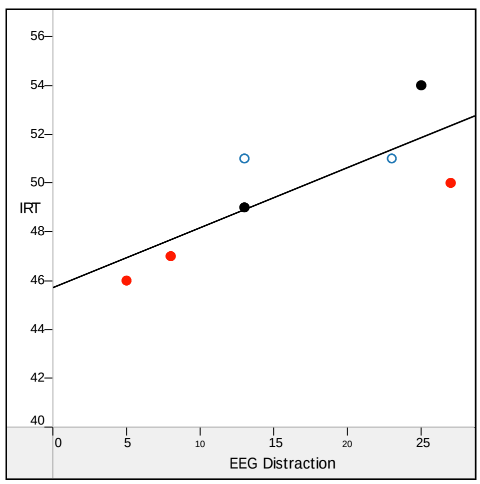

Figure 4. Correlations between EEG-DT and IMMEX™ case difficulty. The EEG-DT levels of two individuals across six cases of three IMMEX™ problem sets were regressed with the difficulties of the different cases previously modeled by IRT analysis from over 8000 student performances.

3.2 Associating B-Alert EEG Measures with IMMEX™ Problem Solving Events

We then extended these results by examining the changes in the three cognitive measures within individual cases to relate them to problem solving events on a second-by-second basis. In these studies the subjects performed three cases of different IMMEX™ problem sets (*Paul's Pepperoni Pizza Palace, Get Organized and Phyto Phiasco*) simulations. During these sessions a monitor recorded when the subject made a test selection which was used for aligning the EEG data with IMMEX™-related events.

In Figure 5 shows such an analysis for one student on the first and second cases of *Phyto Phiasco*. The limited use of background information by this individual, as illustrated by few tests being ordered in the upper left corner of the performance maps, suggests that he had domain knowledge. From the uniform strategic approach across the three cases (ANN nodes 9,15,9) all in close proximity on the neural network topology in Figure 3, as well as the limited HMM State classification (States 4,4,4) suggests this person also had prior problem solving experience.

On the first performance there were sustained (2-4 seconds) periods of high / maximum EEG-E closely associated with the selection and viewing of new and novel data on the screen and reduced levels during the intervals between tests. On the second performance the EEG-E was lower overall (62% vs. 28%) and this was associated with fewer periods of sustained EEG-E. While halfway through the second performance the EEG-E levels showed sustained elevation, this was associated with the subject beginning to take notes (Note Taking) rather than the viewing of novel data (Figure 5).

By IRT analysis, the second case being performed was significantly more difficult than the case performed first and this was reflected by increased levels of EEG-DT distributed throughout the performance (EEG-DT = 5% on performance 1 vs. 25% for performance 2). During the note taking the EEG-DT levels were reduced while the EEG-DT was high. The EEG-WL was not significantly different across performances and fluctuated throughout both episodes and was not obviously correlated with problem solving events, EEG-E or EEG-DT.

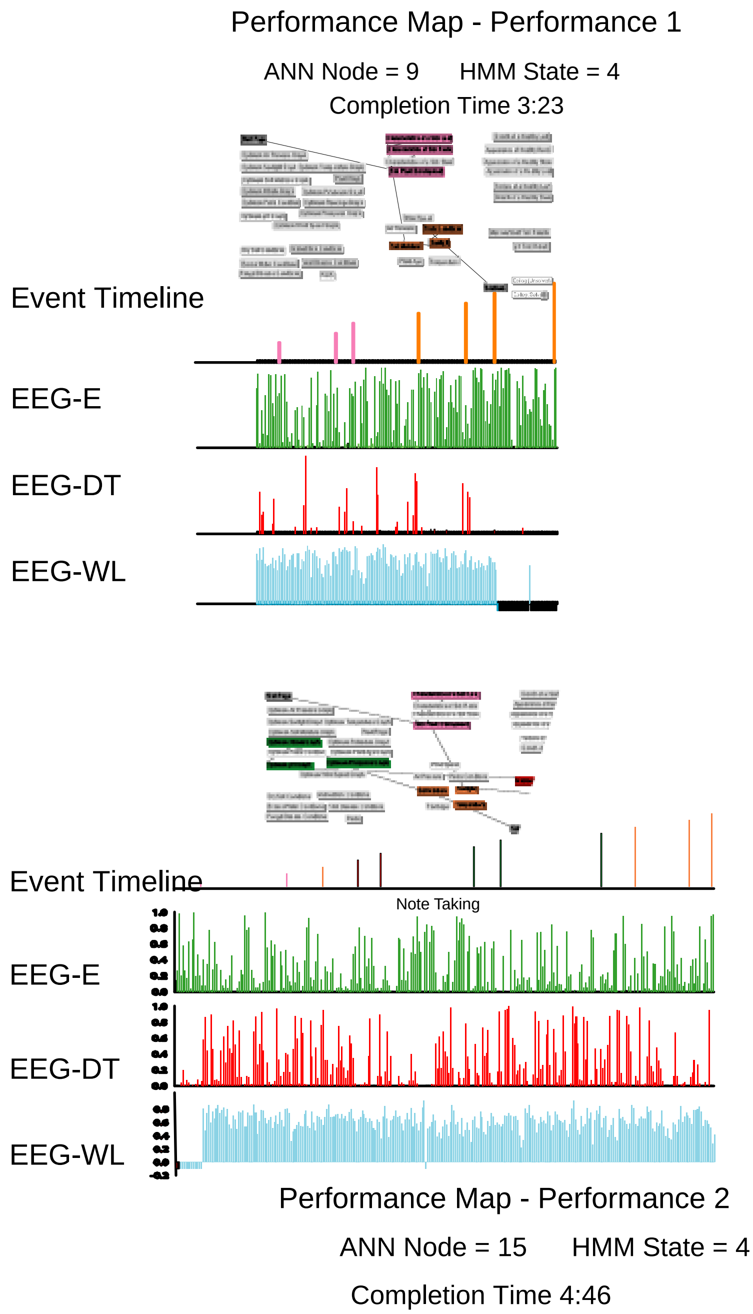

Figure 5. Alignment of EEG-E, EEG-DT and EEG-WL with Problem Solving Events. The real-time B-Alert EEG metrics were aligned with the monitor-recorded problem solving events for two *Phyto Phiasco* performances of one

subject. The upper section of each panel maps how the subject navigated the problem space each time (Stevens et al, 2005). The Event Timeline graphic identifies the sequential test selections for the two performances. They are recorded as 1 for the first test selected, 2 for the second, etc. to make it easier to associate them with the tests in the path maps. The two sets of EEG measures are normalized to the time taken across each performance. The notation 'Note Taking' on the second performance identifies where the subject began taking notes.

IMMEX™ tasks are more open-ended in that students may complete the problem using very few tests or many tests. Most often individuals will conduct testing with different sequences of tests and with varying intervals between tests. As such, it is likely that both the description and definition of the terms Distraction, Engagement and Working Load will change as the B-Alert acquisition and analysis of data is expanded to other complex scenarios. In particular, our previous studies have suggested that EEG-DT may be a heterogeneous and interesting metric within the context of IMMEX™ problem solving. One subject performing *Phyto Phiasco* inadvertently provided the opportunity to begin to explore this metric in more detail (Figure 6).

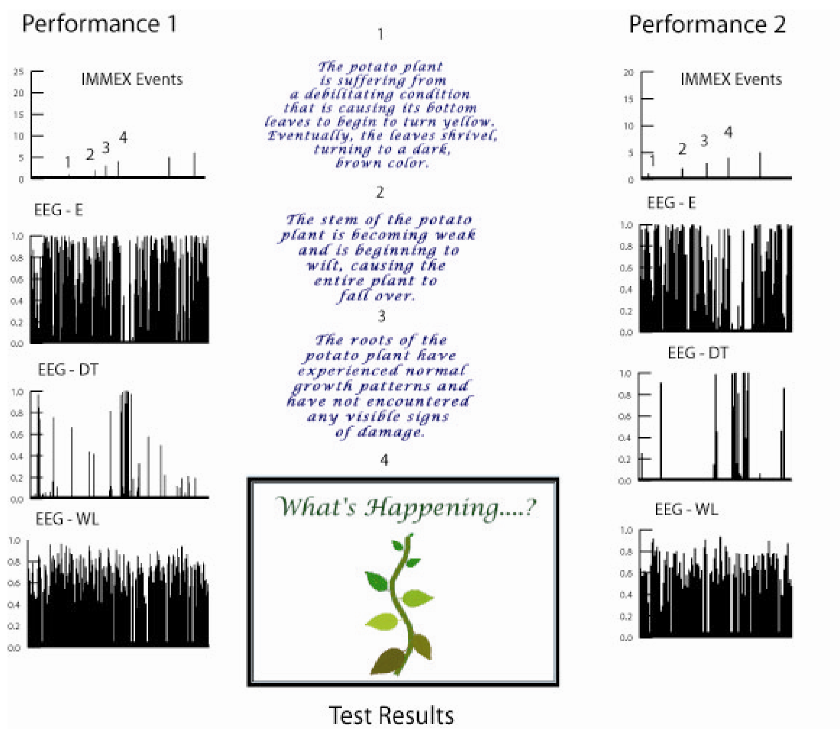

Figure 6 Effects of Text vs. Animated Data Presentation on EEG Metrics. One subject began two cases of *Phyto Phiasco* with the same starting four tests (labelled 1-4). The data in tests 1-3 was presented in a text form and combined presented the same case-related information as did the single fourth test which was an animation. EEG-E, EEG-DT and EEG-WL are presented for case 1 (left) or case 2 (right) beneath the event log for the IMMEX™ performances.

This person began both the first and second cases by examining test data for the Leaves of the Sick Plant, Stems of the Sick Plant and Roots of the Sick Plant all of which presented the information in text form. For the fourth test item an animation of the plant growing and dying was selected which essentially combined the data of the first three tests into a single test. On both performances, there was a high and sustained rise in EEG-DT that accounted for most of the EEG-DT of the total performance. By examining the 1Hz bins of the F_zPO_z and C_zPO_z channels, elevations were seen in the 11, 14 and 33 Hz channels. As no new data / information was supplied by the fourth test, it would suggest that the nature of the presentation may have precipitated significant EEG-DT changes. The participant reported being frustrated by the animated presentation in part due to the redundant information but also because she preferred the verbal presentation mode.

4 DISCUSSION

In this manuscript we have begun to associate EEG correlates of attention and memory with probabilistic models of complex scientific problem solving by integrating the metrics of EEG-DT, EEG-E and EEG-WL with problem solving activities. The original derivation of EEG-DT was through the optimization of train-to-criteria performances from documented cognitive working memory tasks, as well as in defense-related applications. In a study of 9 Navy fleet operators performing a 32 missile salvo on a Tactical Tomahawk Weapons Control System the correlation between EEG measures of workload, engagement and distraction and an expert's observed ratings of attention, stress, confusion and frustration were evaluated (Poythress, 2006). Significant positive correlations ($p<0.05$) were obtained between EEG-distraction and expert observed ratings of frustration, confusion, stress. These data in combination with previous results showing a positive relationship between number of errors and levels of EEG-distraction (Berka, 2006) suggest that this measure may be characteristic of multiple cognitive states including distraction, boredom, confusion and frustration. The investigators plan to explore the possibility of creating several sub-classes of what is currently the EEG-distraction metric.

IMMEX™ tasks are more open-ended in that students may complete the problem using very few tests or many tests. Most often individuals will conduct testing with different sequences of tests and with varying intervals between tests. As such, it is likely that both the description and definition of the terms Distraction, Engagement and Working Load will change as the B-Alert acquisition and analysis of data is expanded to other complex scenarios. In particular, our results suggest that EEG-DT may be a heterogeneous and interesting metric within the context of IMMEX™ problem solving.

One component would be the physical distraction caused by background environmental noises. A more problem-solving related component would be the distraction associated with a student missing the solution to the problem. This form of EEG-DT was often followed by a period of increased engagement. Finally, the correlation between the problem difficulty and EEG-DT measures may suggest an involvement in a more subtle form of problem solving uncertainty. In this regard it may be similar to the metrics being derived by DuRousseau et al, (2005) that relate to one's confidence in an answer prior to submitting it. This correlation was mainly observed in more experienced users where strategies were more limited and the problems were being solved decreasing the other forms of distraction.. As more data is collected, these differences in EEG-DT may provide an opportunity to extend and separate existing measures to reflect these components.

The changes in EEG-E across case performances that we observed were more like what would be expected reflecting novelty of the data presentation on the screen. As EEG-E values generally decreased with experience it would appear that this metric is responding more to the appearance of the data presentation rather than the novelty of the data within the display.

The relative uniform values of EEG-WL across different cases suggest this metric may be quite heterogeneous, or the changes may be quite small and subtle in the context of IMMEX™ problem solving. By beginning to segment the performance into intervals between the different problem solving events such differences were observed and future studies will refine this segment further by examining 5 and 10-second intervals before and after each event.

As increasing numbers of performances are collected and as the models of EEG metrics are refined and aggregated at each ANN node of the problem space topology we will then begin exploring ways of extending existing skill acquisition models through the combination of the two approaches.

Supported in part by National Science Foundation grants REC 02131995, HRD 0429156 and DUE 0126050.

5 REFERENCES

Anderson, J. R. Acquisition of cognitive skill. Psychological Review, 89, 369-406, 1982.

Anderson, J. R. Cognitive psychology and its implications (4th ed.). New York: Freeman. 1995.

Arroyo, I., Beal, C. R., Murray, T., Walles, R., Woolf, B. P. (2004). Web-Based Intelligent Multimedia Tutoring for High Stakes Achievement Tests. 468-477. James C. Lester, Rosa Maria Vicari, Fábio Paraguaçu (Eds.): Intelligent Tutoring Systems, 7th International Conference, ITS 2004, Maceiò, Alagoas, Brazil, Proceedings. Lecture Notes in Computer Science 3220. Springer 2004.

Beal, C. R. (2004). Adaptive user displays for intelligent tutoring software. Cyberpsychology and Behavior, 7, 689-694.

Berka, C., Levendowski, DJ., Cvetinovic, MM., Davis, GF., Lumicao, MN., Popovic, MV., Zivkovic, VT., Olmstead, RE., Real-Time Analysis of EEG Indices of Alertness, Congnition and Memory Acquired with a Wireless EEG Headset, International Journal of Human-Computer Interaction, 17(2), 151-170, 2004.

Berka, C., Levendowski, DJ., Ramsey, C..K., Davis, G., Lumicao, M.N., Stanney, K., Reeves, L. Harkness, Regli,, S., Tremoulet, P.D., Stibler, K., Evaluation of an EEG-Workload Model in an Aegis Simulation, Biomonitoring for Phisological and Cognitive Performance during Military Operations, edited by John Caldwell, Nancy Jo Wesentsten, Proceddings of SPIE Volume 5797, 90-99, 2005.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, Al, Davis, G., Zivkovic, V. T., Olmstead, R.E., Tremoulet, P. D., Craven, P. L. EEG correlates of task engagement and mental workload in vigilance, learning and memory tasks. Journal of Aviation, Space and Environmental Medicine, Special Supplement, 2006, In press.

Cooper, M.M., Cox, C.T. (Jr), Nammouz, M., and Stevens, R. (Submitted). Improving Problem Solving with Collaborative Learning.

DuRousseau, D.R. Mannucci, M. A., and Stanley, J.P. (2005). Will Augmented Cognition Improve Training Results? *in* Foundations of Augmented Cognition v.II. D. Schmorrow (ed). Lawrence Erlbaum Associates, Mahwah, NJ

Ericsson, K. A. (2004). Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. Academic Medicine. 79 (10), S70-S81.

Fabiani, M., Gratton, G. and Coles, M. G. (2000). Event-related Brain Potentials. Handbook of Psychophysiology. J. T. Cacioppo, L. G. Tassinary and G. G. Berntson. Cambridge, Cambridge University Press: 53-84.

Haider, H., and Frensch, P.A. (1996). The role of information reduction in skill acquisition. Cognitive Psychology 30: 304-337

Kohonen, T. (2001). Self organizing maps. (3rd ed.) Springer Series in Information Sciences, Vol. 30, Springer Heidelberg, Germany.

Poythress, M., Russell, C., Siegel, S., Tremoulet, P. D., Craven, P. L., Berka, C., Levendowski, D.J., Chang, D., Baskin, A., Champney, R., Hale, K., Milham, L. Correlation between Expected Workload and EEG Indices of Cognitive Workload and Task Engagement. In Proceedings of 2nd Annual Augmented Cognition International Conference, San Francisco, CA. (In press)

Schneider, W. and Shiffrin, R. M. (1977). "Controlled and automatic human information processing I: Detection, search, and attention." Psychological Reviews 84: 1-66.

Schunn, C. D., & Anderson, J. R. (1999). The Generality/Specificity of Expertise in Scientific Reasoning. Cognitive Science. 13, 337-370.

Schunn, C.D., & Reder, L.M. (2001). Another source of individual differences: Strategy adaptivity to changing rates of success. *Journal of Experimental Psychology: General*, 130, 59–76.

Stevens, R., & Casillas, A. (2005). Artificial Neural Networks. In R. E. Mislevy, D. M. Williamson, & I. Bejar (Eds.), Automated Scoring. Lawrence Erlbaum. pp 259-312.

Stevens, R., & Casillas, A. (2006). Artificial Neural Networks. In R. E. Mislevy, D. M. Williamson, & I. Bejar (Eds.), *Automated Scoring of Complex Tasks in Computer Based Testing: An Introduction.* Lawrence Erlbaum, Mahwah, NJ.

Stevens, R., Johnson, D. F., & Soller, A. (2005). Probabilities and Predictions: Modeling the Development of Scientific Competence. Cell Biology Education 4(1): 42–57. The American Society for Cell Biology.

Stevens, R., Soller, A., Cooper, M., and Sprang, M. (2004) Modeling the Development of Problem Solving Skills in Chemistry with a Web-Based Tutor. Intelligent Tutoring Systems. J. C. Lester, R. M. Vicari, & F. Paraguaca (Eds). Springer-Verlag Berlin Heidelberg, Germany. 7th International Conference Proceedings (pp. 580-591).

Stevens, R., Wang, P., Lopo, A. (1996). Artificial neural networks can distinguish novice and expert strategies during complex problem solving. JAMIA vol. 3 Number 2 p 131-138

Stevens, R.H., & Najafi, K. (1993). Artificial Neural Networks as Adjuncts for Assessing Medical Students' Problem-solving Performances on Computer-based Simulations. Computers and Biomedical Research, 26, 172-187.

VanLehn, K., (1996). Cognitive Skill Acquisition. Annual Review. Psychology. 47: 513-539