Exploring Neural Trajectories of Scientific Problem Solving Skill Acquisition

Ronald H. Stevens[1], Trysha Galloway[1], and Chris Berka[2]

[1] UCLA IMMEX Project, 5601 W. Slauson Ave. #255, Culver City, CA 90230
immex_ron@hotmail.com,
tryshag@gmail.com
[2] Advanced Brain Monitoring, Inc, Carlsbad, CA 92008
chris@b-alert.com

Abstract. We have modeled changes in electroencephalography (EEG) - derived measures of cognitive workload, engagement, and distraction as individuals developed and refined their problem solving skills in science. Subjects performing a series of problem solving simulations showed decreases in the times needed to solve the problems; however, metrics of high cognitive workload and high engagement remained the same. When these indices were measured within the navigation, decision, and display events in the simulations, significant differences in workload and engagement were often observed. In addition, differences in these event categories were also often observed across a series of the tasks, and were variable across individuals. These preliminary studies suggest that the development of EEG-derived models of the dynamic changes in cognitive indices of workload, distraction and engagement may be an important tool for understanding the development of problem solving skills in secondary school students.

Keywords: EEG, Problem solving, Skill Acquisition, Cognitive Workload.

1 Introduction

Skill development occurs in stages that are characterized by changes in the time and mental effort required to exercise the skill (Anderson, 1982, 1995, Schneider and Shiffrin, 1977). Given the complexities of skill acquisition it is not surprising that a variety of approaches have been used to model the process. For instance, some researchers have used machine learning tools to refine models of skill acquisition and learning behaviors in science and mathematics. Such systems rely on learner models that continually provide 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), or identifying when students may require an educational intervention, they still rely on relatively impoverished input due to the limited range of learner actions that can be detected by the tutoring system (e.g., menu choices, mouse clicks).

D.D. Schmorrow, L.M. Reeves (Eds.): Augmented Cognition, HCII 2007, LNAI 4565, pp. 400–408, 2007.

There is a large and growing literature on the EEG correlates of attention, memory, and perception (Fabiani, 2001, Smith, 1999, Berka, 2004, Berka 2006). However, EEG researchers have generally elected to employ study protocols that utilize training-to-criterion to minimize variability across subjects and ensure stable EEG parameters could be characterized. In most studies, the EEG data is not even acquired during the training process, leaving an untapped and potentially rich data source relating to skill acquisition.

While advanced EEG monitoring is becoming more common in high workload / high stress professions (such as tactical command, air traffic controllers), the ideas have not been comprehensively applied to real-world educational settings due to multiple challenges. Some of these challenges are: 1) 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 domain understanding; 2) 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 skill and 3) the strategic variability makes analyzing the patterns of students' problem solving record too complicated, costly, and time consuming to be performed routinely by instructors; nevertheless, there are many aspects of science education that could benefit from deriving data from advanced monitoring devices and combining them with real-time computational models of the tasks and associated outcomes conditions.

This manuscript describes a beginning synthesis of 1) a probabilistic modeling approach where detailed neural network modeling of problem solving at the population level provides estimates of current and future competence, and, 2) a neurophysiologic approach to skill acquisition where real-time measures of attention, engagement and cognitive work load dynamically contribute estimates of allocation of attention resources and working memory demands as skills are acquired and refined.

2 Methods

The IMMEX™ Problem Solving Environment

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

To illustrate the system, a sample biology task called Phyto Phyasco 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 5 Main Menu items which are used for navigating the problem space, and 38 Sub Menu items describing local weather conditions, soil nutrients, plant appearance, etc. These are decision points, as when the student selects them, s/he confirms that the test was requested and is then presented the data. When students feel they have gathered the information needed to identify the cause they attempt to solve the problem.

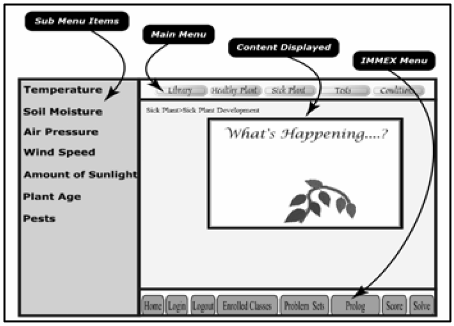


Fig. 1. Sample IMMEX™ simulation. In the Phyto Phyasco simulation, potato plants are dying and the student must identify the cause by examining weather conditions, nutrients, etc. Students navigate throughout the problem space using the Main Menu and select data to make decisions using the Sub Menu Items.

The IMMEX database serializes timestamps of how students use these resources. 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). We use the outputs of this classification to define strategic snapshots of each student's performance. Students often begin by selecting many test items, and consistent with models of skill acquisition (Ericsson, 2004), refine their strategies with time and select fewer tests, eventually stabilizing with an approach that will be used on subsequent problems. As expected, with practice solve rates increase and time on task decreases. The rate of stabilization, and the strategies stabilized with are influenced by gender (Stevens & Soller, 2005), experience (Stevens et al, 2004), and individual or group collaboration.

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 opens a new window to monitor the cognitive demands and the balances of different working memory capacities needed as students gain experience and begin to stabilize their strategies.

The B-Alert® system

A recently developed wireless EEG sensor headset has combined a battery-powered hardware and sensor placement system to provide a lightweight, easy-to-apply method to acquire and analyze six channels of high-quality EEG. 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). Data are sampled at 256 samples/second with a bandpass from 0.5 Hz and 65Hz (at 3dB attenuation). 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. Wavelet analyses are applied to detect excessive muscle activity (EMG) and to identify and decontaminate eye blinks (Berka 2004, Berka 2007).

Subjects and Study
Subjects (n=12) 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 Phyasco, the biology problem described above 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 strategies and cognitive changes across cases as students gained experience. Then 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. For this alignment, we used software (Morea, Techsmith, Inc.) that captures output from the screen, mouse click and keyboard events as well as video and audio output from the users (Figure 2).

The output of the B-Alert software includes EEG metrics (ranging from 0.1-1.0) for distraction (HDT), engagement (HE), and workload (HWL) calculated for each

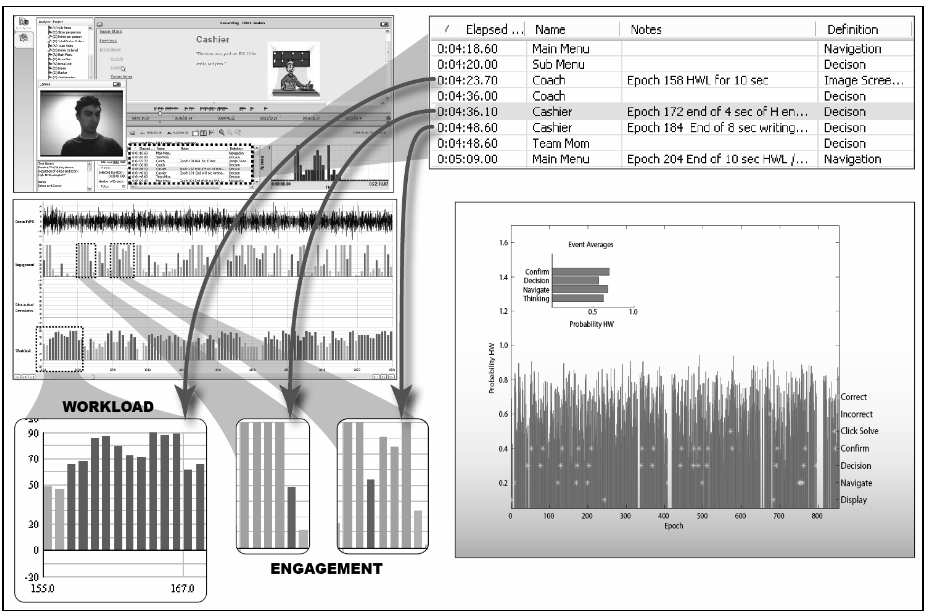


Fig. 2. Relating EEG Workload and Engagement Indexes with Problem Solving Events. The upper left figure is a screen shot of a user (not described in the text) engaged in IMMEX problem solving while keyboard and mouse events are simultaneously recorded; below shows the real-time output of the B-Alert cognitive indexes. Samples of the workload and engagement data streams have been linked with specific events in the log.. In the lower right corner, the timestamps of IMMEX data requests and displays are integrated with the EEG workload indices and then plotted against the one-second epochs of the task. The upper left histograms average the workload indices for each of the IMMEX events including the one second prior to and after the event.

1-second epoch 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 (Berka 2007).

These metrics have proven utility in tracking both phasic and tonic changes in cognitive states, and in predicting errors that result from either fatigue or overload (Berka 2005, Berka 2007). The cognitive indices are expressed as histograms for each 1-second epoch of the problem solving session and show the probability of HWL, HE, or HDT. By integrating the time stamps of data requests with those of the B-Alert system, the navigation, decision, and display-related events can be overlaid onto the cognitive indices.

Increases in Problem Solving Skills Are Not Accompanied by Decreases in Workload or Engagement

We first measured the seconds needed to solve the first, second, and third case of Paul's Pepperoni Pizza (n=7) and calculated the average HWL and HE across the three performances. As shown in Table 1, the time needed to complete the task significantly decreased, however, there were no significant changes in either HWL or HE.

Students Apply Similar Workload to Similar Problems and More Workload to More Difficult Problems

Five of the students also performed 3 cases of Phyto Phyasco which is also a middle school IMMEX problem. There were no significant differences between the HWL (0.64 ± .05 vs. 0.63 ± .05, $p = .42$) and HE (0.51 ± .07, 0.51 ± .04, $p = .92$) across the two problem sets. Two individuals also solved the more difficult high school chemistry problem Hazmat. For both of these individuals the HWL was significantly greater for the three cases of Hazmat than for Paul's Pepperoni Pizza. (Subject 103: 0.76 ± .02 vs. 0.71 ± .03, $p < 0.001$; Subject 247: 0.57 ± .02 vs. 0.49 ± .03, $p < 0.005$).

Table 1. Changes in Time on Task, HWL and HE With Problem Solving Experience

Performance	Speed (seconds)	HWL	HE
1	422 ± 234	.629 ± .07	.486 ± .09
2	241 ± 126	.625 ± .08	.469 ± .08
3	136 ± 34	.648 ± .06	.468 ± .09

The Navigation and Decision-related Events in IMMEX May Be Behaviorally Relevant

For the initial studies we divided the performance into segments related to problem framing, test selections, confirmation events where the student decides whether or not to pay the fee for the data, and closure where the student decides on the problem solution. The examples for this section were derived from one student who was performing the IMMEX mathematics problem Paul's Pepperoni Pizza. The student missed solving the first case, correctly solved the second case, and then missed the third case indicating that an effective strategy had not yet been formulated.

The problem framing event was defined as the period from when the Prolog first appeared on the screen until the first piece of data information is chosen. For this subject the HWL decreased from the first to the third performance (.72 ± .11 vs. .57 ± .19, t = 28.7, p < .001), and engagement increased .31 ± .30 vs. .49 ±.37 t = 4.3, p <.001). The decreased workload was similar to that observed in other subjects; the increasing HE may relate more to the student missing the problem. During the decision-making process, students often demonstrated a cycling of the B-Alert cognitive indexes characterized by relatively high workload and low engagement which then switched to lower workload and higher engagement (Figure 3). The cycle switches were often, but not always associated with selection of new data.

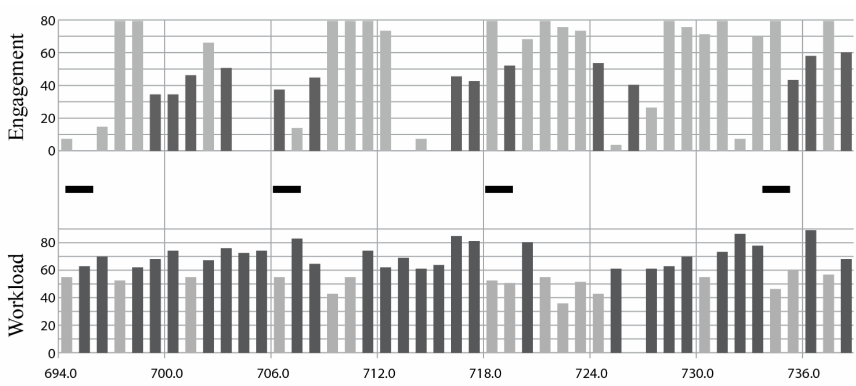
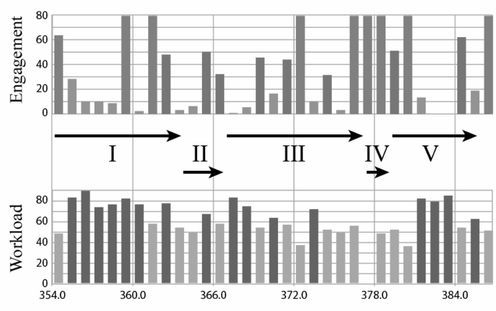


Fig. 3. Fluctuations in HWL and HE during Problem Solving. The bar indicates the epochs where the student made a test selection.

The closing sequences of a problem are a complex process where the student first makes an irrevocable decision to attempt a solution. Then, the he must make a selection choice from an extensive list of possible solutions. Finally, they must confirm their choice. After that they receive feedback on their success / failure; the students have two such solution attempts. The dynamics of HWL and HE for one student's first and second solution attempts of Paul's Pepperoni Pizza are shown in Fig. 4.

In the 10 seconds before deciding to solve the problem (epochs 354 – 364 (I)) there was HWL which decreased as the student made his decision (II, III). Two seconds before the student clicked on and confirmed his choice (epoch 377, IV) there was an increase in engagement which was maintained as the student realized that the answer was incorrect (V).

The workload and engagement dynamics were different on the second solution attempt. Here there was less continuous HWL in the 10 seconds leading up to the decision to solve the problem (Epochs 582- 592, (I, II). At epoch 593 the choice to continue was confirmed, and two seconds before making this decision engagement increased and was maintained during the selection and confirmation process. Between epochs 593 and 596 an incorrect answer was chosen and confirmed (III, IV). At epoch 597 the selection was made and the student learned of the incorrect answer (V).



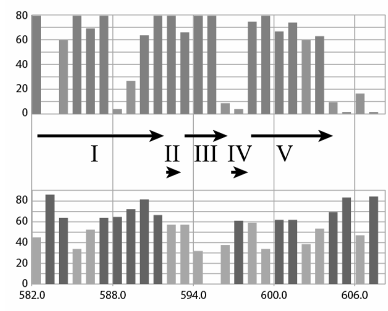


Fig. 4a. Workload and Engagement Events Related to Problem Closure on the First Attempt

Fig. 4b. Workload and Engagement Events Related to Problem Closure on the Second Attempt

Across Performance Changes in HE for Decision-Related Events

Although there are not significant changes in HWL or HE as students develop problem solving skills, we often observed changes in these indices for the Navigation and Decision-related events across performances. For example as shown in Figure 5 (Paul's Pepperoni Pizza), the student correctly solved the first case, had difficulty solving the second case (missing the first try), and then failed to solve the next two cases. The overall workload and engagement levels did not change for this student across the four cases; however, there was a continual increase in the levels of HE for the submenu items where decisions were being made.

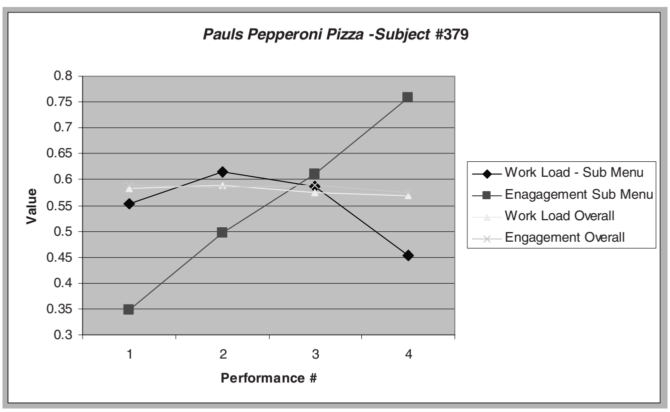


Fig. 5. Changes in Overall HWL and HE and Sub Menu Linked HWL and HE Across Performances

3 Discussion

In this paper we have described a web-based data acquisition architecture and event interleaving process that allows us to map EEG-derived cognitive indices to behaviorally relevant aspects of the students problem solving. An unusual feature of these studies was the application of these technologies to every-day classroom activities that are quite distinct from highly controlled laboratory tasks.

Given the anticipated differences between individual students experience and knowledge, we have focused our initial studies on comparing differences within individuals as skills are developed, rather than compare across individuals. As expected, HWL increased when students were presented with problem sets of greater difficulty. Less expected, however, was the finding that as skills increased, the levels of HWL did not decrease accordingly; suggesting significant mental commitment may be involved during strategic refinement.

By focusing the analyses around relevant problem solving events such as menu navigation and decision making, the changing dynamics of cognitive workload and engagement could be identified. By recording videos of the problem solving process and the user on a second by second basis and interleaving them with EEG cognitive indices through log files generated by IMMEX, the majority of the HWL and HE fluctuations could be linked to observable events such as decision-making and note-taking. Initial studies suggest that decreasing workload and increasing engagement at different events of the problem solving process, such as problem framing and closure, may indicate the student experiencing difficulties and suggest bottlenecks in the learning process. These studies indicate the development of EEG-derived models of the dynamic changes in cognitive indices of workload, distraction, and engagement could be an important tool for understanding the development of problem solving skills in secondary school and university students. Long-term, such models may help target interventions to specific aspects of problem solving where the mental states of an individual reveal barriers to acquiring problem solving skills.

Acknowledgements. Supported in part by grants from the National Science Foundation (NSF-ROLE 0231995, DUE Award 0126050, ESE 9453918) and the U.S. Department of Education (R305H050052).

References

1. Anderson, J.R.: Acquisition of cognitive skill. Psychological Review 89, 369–406 (1982)
2. Berka, C., Levendowski, D.J., Cvetinovic, M.M., Davis, G.F., Lumicao, M.N., Popovic, M.V., Zivkovic, V.T., Olmstead, R.E.: Real-Time Analysis of EEG Indices of Alertness, Cognition and Memory Acquired with a Wireless EEG Headset. International Journal of Human-Computer Interaction 17(2), 151–170 (2004)
3. Berka, C., Levendowski, D.J., Ramsey, C., Davis, K., Lumicao, G., Stanney, M.N., Reeves, K., Harkness, L., Regli, S., Tremoulet, P.D., Stibler, K.: Evaluation of an EEG-Workload Model in an Aegis Simulation, Biomonitoring for Physiological and Cognitive Performance during Military Operations. In: Caldwell, J., Jo Wesentsten, N. Proceddings of SPIE, vol. 5797, pp. 90–99 (2005)
4. Berka, C., Levendowski, D.J., Lumicao, M., Yau, A., Davis, G., Zivkovic, V.: EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning and Memory Tasks. Aviation Space and Environmental Medicine. 78(5, Section II, Suppl.) (May 2007) (in press)

5. Ericsson, K.A.: Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. Academic Medicine 79(10), S70–S81 (2004)
6. Fabiani, M., Gratton, G., Coles, M.G.: Event-related Brain Potentials. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (eds.) Handbook of Psychophysiology, pp. 53–84. Cambridge University Press, Cambridge (2000)
7. Haider, H., Frensch, P.A.: The role of information reduction in skill acquisition. Cognitive Psychology 30, 304–337 (1996)
8. 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)
9. Schneider, W., Shiffrin, R.M.: Controlled and automatic human information processing I: Detection, search, and attention. Psychological Reviews 84, 1–66 (1977)
10. Stevens, R., Casillas, A.: Artificial Neural Networks. In: Mislevy, R.E., Williamson, D.M., Bejar, I. (eds.) Automated Scoring of Complex Tasks in Computer Based Testing: An Introduction, Lawrence Erlbaum, Mahwah (2006)
11. Stevens, R., Johnson, D.F., Soller, A.: Probabilities and Predictions: Modeling the Development of Scientific Competence. Cell Biology Education 4(1), 42–57 (2005)
12. Stevens, R., Soller, A., Cooper, M., Sprang, M.: Modeling the Development of Problem Solving Skills in Chemistry with a Web-Based Tutor. In: Lester, J.C., Vicari, R.M., Paraguaçu, F. (eds.) ITS 2004. LNCS, vol. 3220, pp. 580–591. Springer, Heidelberg (2004)
13. Stevens, R., Wang, P., Lopo, A.: Artificial neural networks can distinguish novice and expert strategies during complex problem solving. JAMIA 3(2), 131–138 (1996)
14. Stevens, R.H., Najafi, K.: Artificial Neural Networks as Adjuncts for Assessing Medical Students' Problem-solving Performances on Computer-based Simulations. Computers and Biomedical Research 26, 172–187 (1993)
15. VanLehn, K.: Cognitive Skill Acquisition. Annual Review. Psychology. 47, 513–539 (1996)