

Using Learner Data to Influence Performance during Adaptive Tutoring Experiences

Robert A. Sottolare, Ph.D.

U.S. Army Research Laboratory
robert.sottolare@us.army.mil

Abstract. During computer-based tutoring sessions, Intelligent Tutoring Systems (ITSs) adapt planning and manage real-time instructional decisions. The link between learner data and enhanced performance is the adaptive tutoring learning effect chain through which learner data informs learner state classification which in turn informs optimal instructional decisions to enhance performance. This paper examines the roles and influence of learner data in both short-term (also called run-time or session) and long-term (also called persistent) learner models used to support adaptive tutoring decisions within the Generalized Intelligent Framework for Tutoring (GIFT). To enhance the usability of tutoring systems and learner performance, recommendations for the design of future learner models are also presented.

Keywords: adaptive tutoring, learner modeling, Intelligent Tutoring Systems

1 Introduction to the Adaptive Tutoring Learning Effect Chain

The adaptive tutoring learning effect chain [1, 2, 3] has evolved over the last three years to define the relationship between learner data, learner states, instructional decisions by the Intelligent Tutoring System (ITS) and learning outcomes (e.g., knowledge acquisition, skill acquisition, performance, and retention). “Instructional decisions” are often used interchangeably with the term instructional strategies/tactics, pedagogical strategies, or instructional principles [4]. For our purposes, we consider strategies separately from tactics. Strategies may be thought of as planning or re-planning by the ITS and include the potential options available to guide the learner toward the achievement of learning objectives while optimizing efficiency (e.g., time on task) and retention (e.g., achieving deep learning).

As shown in Figure 1, the adaptive tutoring learning effect chain [3] identifies that learner states inform instructional strategy selection, and instructional strategies along with instructional context are needed to select optimal instructional tactics (actions by the tutor). Optimal strategies/tactics are measured by comparing the effect size of learning gains (performance, knowledge & skill acquisition, retention) between instructional methods for any given set of learner states and instructional context.

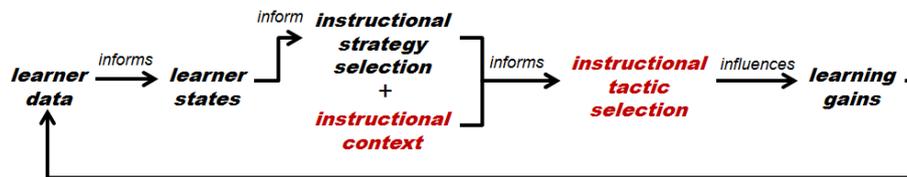


Fig. 1. Adaptive Tutoring Learning Effect Chain [3]

This paper examines the influence of various types of learner data on both pre-exercise and real-time instructional decisions by ITSs with the goal of generalizing design principles for future tutoring systems. As background, we review the relationships between elements within the adaptive tutoring learning effect chain, which is the primary process model within GIFT.

1.1 Learner Data in the Adaptive Tutoring Learning Effect Chain

Learner data is any historical, behavioral, or physiological data used by ITSs to support situational understanding of the learner’s habits, trends, preferences or states. Historical data may be read in from a Learning Record Store (LRS) and used in making both pre-instructional and real-time instructional decisions. The Advanced Distributed Learning Initiative defines a LRS as “a system that stores the tracking statements communicated through the Experience API (xAPI).” Behavioral and physiological data is usually acquired in real-time and used by machine learning algorithms to ascertain current states or project future states.

The ability of ITSs to adapt to the specific needs of individual learners is highly dependent on access to information about the learner, also known as learner data, which can be interpreted through machine learning techniques to classify/predict learner states. Learner data includes real-time data collected during an ITS tutoring session and historical data which may be stored offline in a LRS. The real-time data may include learner decisions, actions, behaviors (e.g., gestures or verbal), or physiological data collected from sensors or other sources. The historical data represents the learner’s most current learning-critical states (e.g., domain competence), descriptive statistics about the learner (e.g., gender), and their long-term traits (e.g., personality or values).

1.2 Learner Modeling in the Adaptive Tutoring Learning Effect Chain

As noted, learner models are constructed from learner data. Models may be descriptive of the learner’s cognitive, affective or psychomotor states. Each of these states may be assessed independently of the instructional context. Performance states may also be assessed based on learner data, but are context dependent and are as-

sessed in real-time. A variety of machine learning methods may be used to evaluate the relationship of learner data to learner states. Typically, learner states are assessed during tutoring sessions. Recently, the U.S. Army Research Laboratory (ARL) evaluated and developed concepts for two types of learner models: short-term (also known as real-time or session learner models) and long-term (also known as persistent learner models). The focus of this paper is on models driven by learner data derived from historical records, sensors (behavioral and physiological data), and interaction with a computer-based tutor (e.g., keystrokes, voice interaction, mouse movement). For example, a learner may be classified in a “super-state” composed of different, but not mutually-exclusive sub-states as noted in Equation (1).

$$\text{Learner State} = f(\text{cognitive, affective, performance, competency}) \quad (1)$$

Cognitive states related to effective tutoring include attention, engagement, and perseverance (also known as grit). Affective states include from least persistent to most persistent: emotions, mood, and personality traits. Performance states are a measure of achievement based on the application of knowledge and skill to a particular task, under a set of specific conditions. Performance may be assessed by measures of speed, accuracy and endurance. Finally, expectations of success may be gauged by domain competency. Competency differs from performance in that competency is potential for success and performance is an actual measure of achievement. For example, based on previous experiences and training, it might be assumed that a student has a high level of expertise in algebra, but this is not a guarantee of success in the future as spacing and repetition might influence the student’s retention and thereby their probability of future success.

1.3 Instructional Context, Strategies, and Tactics in the Adaptive Tutoring Learning Effect Chain

Learner states are used to select strategies (e.g., plans or recommendations) and tactics (actions). Within the architecture of the Generalized Intelligent Framework for Tutoring (GIFT) [5], strategies are domain-independent and tactics are domain-dependent. The more the tutor knows about the learner’s states, the more effective the strategies and tactics will be. Strategies are selected to narrow the tutor’s options for tactics and are based solely on the learner’s states (e.g., cognitive, affective, and performance) within GIFT.

It is important to note that instructional strategies may be either macro-adaptive or micro-adaptive. *Macro-adaptive strategies* are employed by ITSs which have the capability to retrieve and use information about the learner prior to the tutoring session. This information is generally stored offline in a LRS as part of the long-term learner models (LTLM) and used to initialize the short-term learner models (STLM). Macro-adaptive strategies are used to plan a list of available tutoring sessions, lessons, or courses based on the learner’s domain competency and other factors.

Micro-adaptive strategies are used to plan near-term actions during tutoring sessions based on the learner evolving states. For example, a micro-adaptive strategy could be error-sensitive feedback where an ITS evaluates answers to problems and initiates a recommendation for a reflective prompt when the learner misses three problems in a row. Both macro-adaptive and micro-adaptive strategies require instructional context to move forward with optimal tactic selection.

Instructional context refers to the instructional setting and environment (e.g., classroom, at home or in transit) in which the learner is receiving tutoring as well as any computer-based environments (e.g., game-based tutoring; mobile learning). For narrative-based tutoring, understanding who, what, where, and when within a narrative (storyline) aids in describing the instructional context and the tactical options available to the ITS. Context might also be described in terms of learning objectives and concepts to be learned. Since tactics are domain-dependent, instructional context comes into play when selecting tactics.

Instructional tactics are the actions taken by the ITS from a set of available options. Actions are taken in response to strategies (plans or recommendations) selected and the context of the instruction.

1.4 Desired Outcomes in the Adaptive Tutoring Learning Effect Chain

Finally, desired outcomes include *learning outcomes* and may include other desirable traits developed as a result of the adaptive tutoring process. Learning outcomes refer to measurable differences in the effect size of performance, knowledge acquisition, skill acquisition or retention. If the methods employed are accurate in assessing the learner's state, and selecting optimal strategies and tactics to reinforce deep learning, then we can expect improvements in the learning outcomes.

In general, ITSs are concerned with learning outcomes which include knowledge and skill acquisition (also known as learning), performance (the application of knowledge and skill) and retention (the ability of the learner to recall and use knowledge and skill in the future). However, ARL has recently begun to evaluate the efficacy of assessing and influencing other outcomes related to desirable attributes which include, but are not limited to: perseverance (grit), adaptability, cooperation, creativity, and critical thinking. These desirable traits may directly contribute to the effect of adaptive tutoring on learning.

1.5 Adaptive Tutoring within GIFT

GIFT is an open-source, modular, service-oriented tutoring architecture which supports authoring of ITSs, development of automated instructional methods, and analysis of ITS technologies [5]. In real-time (during instruction), GIFT (see **Fig. 2**) is an instance of the adaptive tutoring learning effect chain previously described in

this paper. The learner model in GIFT captures and interprets learner data to predict learner states. A portion of this learner data is targeted to be historical data from the LTLM within a LRS. Learner states are used by the pedagogical module in GIFT to generate domain-independent recommendations based on learner states (performance, engagement, and affect). These strategy recommendations are used by the domain module to select the tactics which are presented by the tutor. At the end of the ITS session, competency states, trends, and newly identified traits may be written to the LTLM.

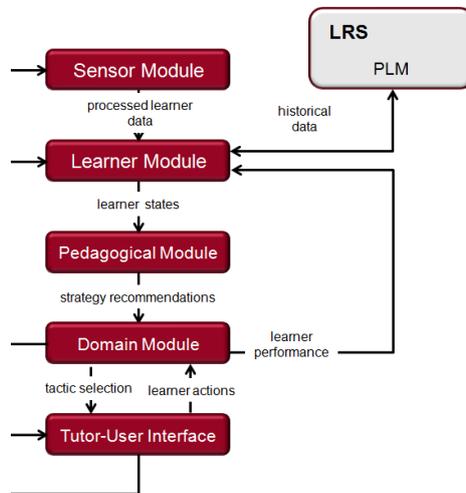


Fig. 2. LTLM in Relationship to GIFT

2 The Influence of Learner Data in Instructional Decisions

This section discusses how long-term and short-term learner data might influence instructional decisions by ITSs. Specific variable of interest (learner states, traits, and data) are also reviewed in terms of their potential inclusion in long-term and short-term learner models.

2.1 Long-term Learner Data for Tutoring Decisions

As discussed previously in this paper, ARL has begun to evaluate concepts for the LTLMs. It is anticipated that LTLMs will contain achievement data from ITS sessions, but will not retain all learner data from each ITS session. For example, sensor data collected during recent experiments conducted by ARL tally several gigabytes and while it might be useful to retain this data for analysis into the future, it is more useful to retain simple descriptors of states and achievements in the LTLM which may be used later to initialize the STLM within the ITS just prior to runtime.

Learner data of interest for a LTLM which have high potential to influence real-time instructional decisions include: learner demographics, personal and cultural values, personality, preferences, interests, domain competencies, and long-term goals. Each of these LTLM variables of interest has the potential to influence learner engagement and motivation by personalizing scenario content and adapting interactions. Domain competency is essential as a measure of expected performance against which actual performance should be constantly assessed.

The LTLM has also been called a “persistent learner model”, but we have identified the LTLM as a dynamic model which changes with each experience and achievement. The LTLM also changes over time. For example, domain competency for marksmanship might be assessed as “expert”, but marksmanship skills erode over time if appropriate refresher training is not undertaken on a regular basis. This should be reflected in a dynamic LTLM and used by ITSs to assess current competency.

2.2 LTLM Variables of Interest

Most learner models today focus on session-level modeling of the learner’s progress within a lesson or course. Often the effect of a tutoring session is captured only in terms of the competency impact as a badge of achievement. Other useful learner information which is captured through assessments or surveys is only used during that session or is totally ignored.

STLMs may be used to populate LTLMs with the most current competency states. LTLMs may also contain other factors, traits, states, and goals which can be used to initialize the STLM and influence the decisions of the tutor. However, most ITSs do not segregate learner models into persistent (long-term) and session (short-term) data. Methods are needed for capturing, storing, retrieving, and using long-term learner data and the potential exists to support more effective ITS instructional decisions by having a long-term model of the learner.

While there are no standards or consensus for what should be in either a STLM or LTLM, we have chosen to focus this paper on evaluating potential variables of interest for inclusion in the LTLM. Specifically, we discuss the influence of these variables in the context of selecting macro-adaptive (pre-instruction) strategies within the adaptive tutoring learning effect chain. While it is desirable to develop ITS strategies that can be generalized across large sections of the population, it is more likely that instructional strategies will be found to generalize within demographics, traits or states. However, these strategies must be shown to be effective in all domains in which they are intended to be used. So what are the factors to be considered in assessing the effectiveness of a given strategy?

Effectiveness may be measured by the ability of each strategy to support the goals identified within each domain. As noted earlier, instructional goals can be directly

tied to specific measures of effect for learning (l), performance (p), retention (r), engagement (e), and motivation (m).

If D is the set of all domains taught by an ITS and s is a specific instructional strategy, then Ds is the set of domains where a specific instructional strategy can be implemented. Since macro-adaptive strategies are implemented during pre-instruction, learner states may not be available, but trait data may be via LTLM. If T is the set of all traits tracked in the LTLM, then Ts is set of traits available when a specific strategy is implemented, and t is a specific trait. The effectiveness of a strategy, s , in each domain may be expressed as a function: $E(s, t, l, p, r, e, m)$ for the strategies under test: demographics, values, personality, preferences, interest, competencies, and long-term goals. Each of these strategies will be described in more detail below.

Demographics.

Demographics refer to descriptors of a group. For example, the following descriptors may be significant in making instructional decisions since they represent groups which may differ in learning strategies and learning styles: gender, culture (e.g., religious or social), age, affiliations, and education. Holden & Sinatra [6] identified demographics as part of long list of domain-independent variables of interest. They also noted that while this type of information is useful, it alone may not be a sufficient basis for adapting tutoring for any individual. This said, demographics may be sufficient for initializing learner models, but empirical evidence is needed to support any micro-adaptive strategy or tactical decisions.

Personal and Cultural Values.

Values are an individual's or group's basis for judging the worth of something. Values include, but are not limited to: honesty, integrity, loyalty, candor, clarity, power, and respect. They drive one's beliefs, attitudes and personal norms, and ultimately behavior including displays of affect (emotion). Personal values are individually defined, and determine what each person believes is good, useful, beautiful or desirable, but they are also influenced by the culture of affiliated groups (e.g., political, geographic, educational, family, religion, or social).

Values can be assessed over time by interaction with an ITS or through the administration of a survey like the Rokeach Value Survey [7]. Values may then be used in an instructional context by the tutor to drive emotions or engagement. This might be useful in assessing decision-making in difficult moral dilemmas or other stressful conditions.

Personality, Preferences and Interests.

A byproduct of our value systems are beliefs about and attitudes toward people, things, methods, and even ideas. This ultimately results in preferences. We take in information about the world and we judge things [8]. The things we “like” are preferences. Preferences manifest themselves in models of personality type [9] or learning styles [10]. Instructional strategies which involve or address preferences may lead to higher levels of engagement and thereby high levels of learning.

Of close relationship to preferences are interests. Interests can be more formal as in professional or educational activities such as science, mathematics, design, or art, or they might be less formal such as hobbies or sports. Interests run the gamut from cognitive (e.g., puzzles) to physical (e.g., golf) to social and collaborative activities. Interest-based instruction correlates with motivation and appears to promote positive short and long-term learning in classrooms [11]. Carnegie Learning has ongoing research to assess the impact of adapting instruction to support individual interests which should confirm the transfer of interest strategies from classrooms to one-to-one ITS sessions. If this connection between interest and learning proves viable, conducting surveys to capture interest measures in the LTLM and then using these to adapt content to fit learner interest may be a practical domain-independent strategy.

Domain Competencies and Long-term Goals.

Finally, we examine the effect of domain competencies and long-term goals on macro-adaptive strategies. Domain competencies are indicators of the ability of a learner to complete a task successfully (effectively and efficiently). It may be considered a measure of the learner’s potential achievement whereas performance is a measure of actual achievement. Domain competency may be critical in developing macro-adaptive strategies related to the presentation of options to the learner. For example, the ITS might develop a hierarchical list of courses or lessons for the learner to consider. As shown in **Fig. 3**, this strategy considers both competence data from the LTLM and may also consider a pre-course assessment as needed.

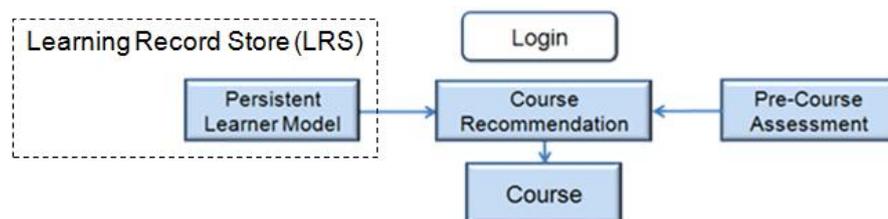


Fig. 3. Macro-Adaptive Strategy to Recommend Courses

2.3 Short-term Learn Data for Tutoring Decisions

During tutoring sessions, the ITS STLM may use data from the LTLM to initialize learner states (e.g., competency). The STLM also uses a constant flow of continuous and discrete-event data to update these learner states. This real-time data is generated by sensors, learner interactions, and learner input in response to content and tactics presented by the tutor. The STLM uses this data to assess transitions between states.

Sensor Data.

The sensor data (behavioral and physiological measures) may be used as input for Dynamic Bayesian Networks, decision trees, production rules, Markov Decision Processes, clustering techniques, or other machine learning techniques to assess learner states. Sensor data may be filtered or thinned to support rapid assessments.

While human tutors use the behaviors of the learner to assess their states (e.g., confusion, frustration), computer-based ITSs have the advantage of being able to also evaluate physiological measures (e.g., heart rate) directly to assess the learner's cognitive, affective, and physical states in near real-time.

A goal is to use sensor data to build generalized models of cognition and affect which can be reused across the user population or at least from day-to-day within the same individual. While affect, specifically mood, has been shown to be a significant predictor of performance [12], preliminary findings from recent experimentation has shown that generalized models of cognition and affect are proving to be impractical as classification models are not reusable [13]. In particular, models which are physiology dependent vary widely based on the learner's stimulant consumption (e.g., caffeine) and sleeping habits. This variation limits the need for interaction between the LTLM and STLM and leaves us with only one option: real-time modeling. Brawner (2013) [7] found that real-time classifiers of affect are of good quality (~80% accurate) while real-time classifiers of cognition are not as good (<60% accurate).

Interaction and Input Data.

Learner interaction through the tutor-user interface (TUI), or other input devices (e.g., natural language interface) provide another method of communication between the learner and the computer-based tutor. The interaction data and pace of input may be an indicator of the learner's confusion, an influential state for learning. The ITS may also present real-time surveys or tests to assess learner progress toward stated learning objectives.

3 Design Considerations for Future Intelligent Tutoring Systems

This section addresses areas of consideration for the design of future ITSs including learner data acquisition and learner modeling to promote more efficient and effective real-time tutoring sessions.

3.1 Considerations for Learner Data Acquisition

There are an increasing number of methods to acquire learner data. The number and types of low-cost commercial sensors are increasing including unobtrusive sensors which can acquire data without disrupting the learning process. Sensors appearing more frequently in the literature due to their low-cost and multi-modal capabilities include Microsoft's Kinect and a variety of webcams. The Kinect provides high frequency data streams which include: posture and facial markers which can be used to interpret engagement and emotional states; gaze direction to assess attention and engagement; and pixel gradation which can be used to assess heart rate at a standoff distance. Depending on software, webcams can also be used to assess behavior and physiology from a distance. The trends leading to more unobtrusive sensing of learner states bodes well for moving ITSs from their traditional static desktop computer mode to a true mobile learning environment.

3.2 Considerations for Learner Modeling

While sensor data may point to a particular learner state other data may point to a conflicting state. For example, sensor data may indicate low engagement, but interaction and performance data could indicate a higher state of engagement. For this reason, it is important to consider agent-based tutoring architectures to assess and deconflict ambiguous learner states. Master agents associated with each learner should be designed to track/deconflict assessments from child agents responsible for appraising single data streams (e.g., sensors). Additional research is needed to weight the assessments of individual agents in terms of their reliability.

3.3 Considerations for Tactic Presentation

Once a decision has been made by the ITS master agent regarding an optimal strategy, a tactic (e.g., change in challenge level or feedback) must be selected based on that strategy and the instructional context. If the tactic selected is feedback, options available for delivery or presentation of the tactic should be considered in the ITS design. Options for tactics presentation generally vary by source. Feedback might be presented by a virtual human, text in a chat box, multi-media presentation or by an undefined source also known as the "voice of God". Feedback may be direct or indirect (e.g., reflective prompts). Consideration should be given to allowing the learner to initiate interaction, ask questions, and demonstrate understanding of concepts.

4 Conclusions

Whether we are discussing human tutors or ITSs, it is evident that making decisions based on learner data has advantages over making decisions in the absence of learner data. The goal for ITSs is to be able to use learner data to assess their states and use this as a basis for formulating instructional strategies and selecting tactics. There are multiple barriers to reaching this goal, but two stand out. One barrier to reaching this goal is the ability to generalize learner assessment techniques across the learner population and even within the same learner across time. A second barrier to reaching this goal is the ability to apply learner assessment techniques somewhat independently of the task/concepts being learned.

One option is to develop real-time techniques to overcome differences in modeling learner states which vary over time. Real-time techniques are more appropriate for physiological learner data sources (e.g., heart rate sensors) which can be affected by things like caffeine intake, sleep quality, and environmental factors (e.g., stress). By assessing learner states in real-time, we eliminate generalized models based on historical individual or population data. Learner models are unique to each person and time, and may be considered perishable and not to be reused. The caveat to this option is competency which is generally based on previous achievements and experiences. Competency is useful in initializing learner models and making course selections.

A second option is to evaluate strategy and tactic selection separately per the Generalized Intelligent Framework for Tutoring (GIFT) [5]. Within GIFT, strategies are domain-independent and selected based only on learner states which include cognitive state, affective state and generic performance (at expectation, below expectation or above expectation). Once a strategy (e.g., prompt for additional information about current concept) is selected, GIFT then matches this strategy with the instructional context (e.g., concept being learned) to support tactic (e.g., action) selection and presentation by the ITS. This model is efficient and allows the ITS to develop strategies (plans) based on current and projected learner states thereby promoting learner-centric strategies for instruction.

5 References

1. Sottolare, R. A. (2012). Considerations in the Development of an Ontology for a Generalized Intelligent Framework for Tutoring. In Proceedings of the *International Defense and Homeland Security Simulation Workshop 2012*, Vienna, Austria.
2. Fletcher, J.D. and Sottolare, R. (2013). Shared Mental Models and Intelligent Tutoring for Teams. In R. Sottolare, A. Graesser, X. Hu, and H. Holden (Eds.) *Design Recommendations for Intelligent Tutoring Systems: Volume I - Learner Modeling*. Army Research Laboratory.
3. Sottolare, R., Ragusa, C., Hoffman, M. & Goldberg, B. (2013). Characterizing an adaptive tutoring learning effect chain for individual and team tutoring. In Proceedings of the

Interservice/Industry Training Simulation & Education Conference, Orlando, Florida, December 2013.

4. Nye, B., Sottolare, R., Ragusa, C., and Hoffman, M. (2014, *in press*). Prologue - Defining Instructional Challenges, Strategies, and Tactics for Adaptive Intelligent Tutoring Systems. In R. Sottolare, A. Graesser, X. Hu, and B. Goldberg (Eds.) *Design Recommendations for Intelligent Tutoring Systems: Volume 2 - Instructional Management*. Army Research Laboratory, Orlando, Florida.
5. Sottolare, R.A., Brawner, K.W., Goldberg, B.S. & Holden, H.K. (2012). The Generalized Intelligent Framework for Tutoring (GIFT). Orlando, FL: U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED).
6. Holden, H. & Sinatra, A. (2013). Chapter 8 – The Need for Empirical Evaluation of Learner Model Elements. In Sottolare, R., Graesser, A., Hu, X., and Holden, H. (Eds.) *Design Recommendations for Intelligent Tutoring Systems: Volume 1- Learner Modeling*. Army Research Laboratory, Orlando, Florida. ISBN 978-0-9893923-0-3.
7. Braithwaite, V. A., & Law, H. G. (1985). Structure of human values: Testing the adequacy of the Rokeach Value Survey. *Journal of personality and social psychology*, 49(1), 250.
8. Jung, C.G. (1971). "Psychological Types". *Collected Works of C.G. Jung*, Volume 6. Princeton University Press. ISBN 0-691-097704.
9. Myers, I.B. & P.B. Myers (1995). *Gifts Differing: Understanding Personality Type*. Mountain View, CA: Davies-Black Publishing. ISBN 0-89106-074-X.
10. Felder, R. and Spurlin, J. (2005). *Applications, Reliability and Validity of the Index of Learning Styles*. TEMPUS Publications, UK.
11. Tomlinson, C. A., et. al. (2003). Differentiating Instruction in Response to Student Readiness, Interest, and Learning Profile in Academically Diverse Classrooms: A Review of Literature. *Journal for the Education of the Gifted* 27, 119-45.
12. Sottolare, R. & Proctor, M. (2012). Passively classifying student mood and performance within intelligent tutoring systems (ITS). *Educational Technology Journal & Society*, 15:2, 104-114.
13. Brawner, K. W. (2013). Modeling Learner Mood In Realtime Through Biosensors For Intelligent Tutoring Improvements, *University of Central Florida*.