Effect of Bayesian Student Modeling on Academic Achievement in Foreign Language Teaching (University Level English Preparatory School Example)*

Burak Galip ASLANa
Izmir Institute of Technology

Özlem ÖZTÜRKb
Dokuz Eylül University

Mustafa Murat İNCEOĞLUC
Ege University

Abstract
Considering the increasing importance of adaptive approaches in CALL systems, this study implemented a machine learning based student modeling middleware with Bayesian networks. The profiling approach of the student modeling system is based on Felder and Silverman’s Learning Styles Model and Felder and Soloman’s Index of Learning Styles Questionnaire. The questionnaire was adapted to Turkish for this experimental study conducted with respect to the visual/verbal and active/reflective dimensions of the model. A topic in EFL was chosen for the learning content design, which was also carried into the digital domain and remastered as separate learning scenes for different learning styles. Computer software was also implemented to carry out the experimental learning processes. A quasi-experimental pre-test, post-test design was conducted with 46 volunteers, with 23 students assigned each to a control and an experimental group to compare academic achievement between student-based learning and conventional computer-based learning. No significant difference was found in academic achievement between the control and experimental groups after the experimental treatment. The diagnostic performance of the proposed student modeling system was also compared with performances from similar studies. This student modeling system had a successful prediction rate of 41% on the visual/verbal dimension and 54% on the active/reflective dimension, respectively.

Key Words
Academic Achievement, Bayesian Networks, Computer Aided Language Learning, Felder and Silverman’s Learning Styles Model, Student Modeling.

* This research is a part of first author’s doctoral thesis.

a Burak Galip ASLAN, Ph.D., is currently working as an instructor in the Computer Engineering Department at Izmir Institute of Technology. His research interests include artificial intelligence and distance education. Correspondence: Izmir Institute of Technology, Department of Computer Engineering, Urla 35430, Izmir, Turkey. Email: bgaslan@ieee.org

b Özlem ÖZTÜRK, Ph.D., is currently working as an instructor in the School of Foreign Languages at Dokuz Eylül University. Contact: Dokuz Eylül University, School of Foreign Languages, Buca 35390, Izmir, Turkey. E-mail: ozturk.ozlem@deu.edu.tr

c Mustafa Murat İNCEOĞLU, Ph.D., is currently a professor of Computer Education and Instructional Technology. Contact: Ege University, Faculty of Education, Department of Computer Education and Instructional Technology, Bornova 35100, Izmir, Turkey. Email: mustafa.inceoglu@ege.edu.tr
The development of advanced communication techniques and widespread use of the Internet makes knowledge acquisition without geographic limitations much easier. These changes have triggered the transformation of student needs, following the path from traditional synchronous education to semi-synchronous and blended e-learning environments that are mostly realized by the clicks of a common mouse on a common software browser. While the teacher’s role shifts from solely unidirectional teaching to facilitating the use of learning materials, the student’s role also shifts from passive learning to taking responsibility for the learning process and continuous self-assessment. As a result, during e-learning, the student usually becomes the decision maker. There are, of course, pros and cons of having such individual responsibility. The positive and negative effects of e-learning on online students have been studied in detail by Cantoni, Cellario, and Porta (2004). One of the most important negative effects of e-learning is the student’s solitude during learning sessions because of limited interaction. In such cases, student dropouts become a common reason for failure when the learning environment does not meet their expectations (Katz, 2002; Levy, 2007), while student-centered learning environments can enhance students’ motivation by using intelligent computer systems.

Even when students and teachers synchronously collaborate for a common purpose, they usually exist in different psychological, environmental, and geographical conditions. When the student’s self-control is in question, it is crucial that the interaction between the student and the learning system must be maintained at a high level. The variety of quite different, individual methods used by students during their learning must not be overlooked when interaction is what matters (Baldwin & Sabry, 2003).

It would not be surprising to think that the underlying component of an e-learning system is more or less sophisticated computer software designed to work with specific hardware requirements. It would then not be surprising that the processing superiority of computers over human beings lies at the heart of student-centered learning in the form of customized service possibilities. There are many customization and personalization systems installed in end-user computer software solutions nowadays. Correspondingly, adaptable and adaptive approaches are frequently used in e-learning systems at the present time (Brusilovsky, 2001). The affirmative influence of organizing learning environments with respect to students’ learning styles is highlighted in many studies (Brown, Cristea, Stewart, & Brailsford, 2005; Liegle & Janicki, 2006; Stash, Cristea, & De Bra, 2004; Sun, Ousmanou, & Williams, 2004). In this study, an adaptive student modeling middleware has been realized and integrated into an e-learning software system. The concept involves recording students’ keyboard and mouse data during learning sessions and then diagnosing the individual student’s learning style with the help of an expert system based on a Bayesian network.

A topic in the EFL curriculum has been chosen as the learning content offered to students. Hence, two different learning contents about “IF-Clauses” were designed, one with learning styles taken into account and the second, without. The proposed adaptive system follows learning content that takes learning styles into account. It should be noted that both learning contents are pedagogically equivalent. Two different measurement tools were used for evaluating the success of this proposed system. First, the 44-item Index of Learning Styles Questionnaire developed by Felder and Soloman (1991), which is based on Felder and Silverman’s Learning Styles Model (Felder & Silverman, 1988), was adapted to Turkish (Atman, İnceoğlu, Öğretmen, & Aslan, 2009). Then a 24-item achievement test on IF-Clauses was prepared.

It should be noted that almost all the studies in the literature focus on diagnosis of learning styles. This study measures academic achievement, which is an important contribution of this paper. Hence, this study has two main goals. The first goal is the evaluation of the proposed student modeling system results with respect to the Turkish-adapted ILS questionnaire results. This study’s learning styles prediction performance is compared to other studies in the literature. The second study goal is utilizing the benefits of learning styles for academic achievement. The hypothesis tested in this study is: “Learning-styles based tutoring results in an increase in academic achievement with respect to traditional web-based tutoring.” The results achieved in regard to both goals are discussed at the end of the paper.

Method

There are three main learning theories, namely the behavioral approach, the cognitive approach and constructivism (Jonassen, 1991). Learning
management system software should ideally be developed with respect to all three of these learning theories. However, the Internet medium introduces a self-driven environment for learners, emphasizing the importance of constructivism (Alonso, Manrique, & Vines, 2009; Hamid, 2002; Huang, Liu, Chu, & Cheng, 2007; Lester, Stone, & Stelling, 1999). Constructivism is based on self-motivated and student-centered environments (Piaget, 1977), where students rely on their own insight and experiences throughout the learning process (Karagiorgi & Symeou, 2005). Hypermedia environments provided by learning management systems introduce many flexible choices and promote student self-motivation based on the constructivist approach (Pear & Crone-Todd, 2002). Learning environments can be organized with respect to learning styles with three main approaches (Edmonds, 1981) namely, adapted systems, adaptable systems, and adaptive systems where student profiles maintaining learning-styles information are the core of such systems (Popescu, Trigano, & Badica, 2007).

Widespread use of web-based learning promotes the use of adaptive systems (Brusilovsky, 2004). As student profiles are subject to change over time (Brusilovsky, 1996), dynamic update of student profiles is an important issue (Triantafillou, Pomportis, & Demetriadis, 2003). The most crucial issue in maintaining student profiles is creation of the student model in adaptive systems (Esposito, Lichelli, & Semeraro, 2004). Considering that the student modeling process is itself a diagnosis problem (Yudelson, Medvedeva, & Crowley, 2008), it is not surprising to see machine learning mechanisms build in student modeling layers of many studies (Aslan & Inceoglu, 2007; Tsiriga & Virvou, 2004). There are four mains problems to be addressed when machine learning is used with a student modeling system: the need for a large amount of data (Beck & Woolf, 2000), computational complexity, the crucial need for tagged data, and the concept drift problem (Webb, Pazzani, & Billsus, 2001).

Measures

A client-server based, browser-shell software was developed in this study to test the proposed student modeling system. The student modeling system focuses on the learning styles of students based on Felder and Silverman's (1988) Learning Styles Model (FSLSM). Since then, Felder and Henriques (1995) have discussed the research potential of the FSLSM model for foreign language teaching; the learning content was chosen to be teaching English as a foreign language, with the limitation of teaching only the IF-Clauses topic because of this study's time and resource constraints. The learning contents of the IF-Clauses topic were developed with respect to chosen learning dimensions of FSLSM. The ILS questionnaire was also adapted to Turkish and a multiple-choice academic achievement test on IF-Clauses was prepared for testing this study's hypothesis. The representation of content is based on an adaptive content distribution approach, which focuses on supplying different learning-style based content to students with the corresponding learning styles.

There are several adaptivity levels in learning management systems, depending on the degree of software based autonomy (Opperman, Rashev, & Kinshuk, 1997), and since adaptive systems (fully automatic) are widely used in recent studies (Brown, Brailsford, Fisher, & Moore, 2009; Brown, Brailsford, Fisher, Moore, & Ashman, 2006) an adaptive system approach has been used in this study. The four main approaches for implementing adaptivity in web-based e-learning systems are: adaptable content selection and navigation based on students, problem solving support, adaptive content presentation, and adaptive navigation support for students (Brusilovsky, 1999). This study is based on adaptive content presentation. In such an adaptive system, either the content itself modifies itself with respect to the student model or different pieces of pre-generated learning content are displayed based on the student model (Papanikolau, Grigoriadou, Magoulas, & Kornilakis, 2002). This study focuses on presentation of pre-generated learning content designed with respect to different learning styles for different student models. The student models can be generated for three different purposes, and different types of data are required (Ragnemalm, 1996): Style-based approaches rely on gathering data about learning styles, motivation-based approaches rely on gathering data to increase student motivation (Martens, Gulikers, & Bastiaens, 2004), and knowledge-based approaches rely on gathering data about students' knowledge levels to set difficulty levels for questions and to select related learning content to compensate for students' weak points. This study is based on capturing learning styles of students with the help of student models. Although there are many learning styles in the literature, many researchers underline the efficiency of Felder and Silverman's Learning Styles Model (FSLSM) in computer-
aided learning systems and adaptive web-based e-learning systems (Carver, Howard, & Lane, 1999; Kuljis & Liu, 2005; Paredes & Rodríguez, 2004; Parvez & Blank, 2007; Sharp, 2004; Stash et al., 2004; Zualkernan, 2007). Researchers also prefer FSLSM for its ease of implementation in online learning and robust learning dimensions (Baldwin & Sabry, 2003), in addition to many studies on the validation and verification analysis of the ILS questionnaire (Cook & Smith, 2006; Felder & Spurlin, 2005; Litzinger, Lee, Wise, & Felder, 2005, 2007; Van Zwanenberg, Wilkinson, & Anderson, 2000; Zywno, 2003). Student modeling process is naturally an intractable diagnosis problem (Millan, Agosta, & Perez de la Cruz, 2001; Self, 1988; Yudelson et al., 2008). Hence machine-learning approaches have been frequently used for diagnosis and prediction of problems in learning management systems (Brna & Cox, 1998; Cumming, 1998). Bayesian networks are also being used to address user modeling problems in different types of systems (Heckerman, Horvitz, & Rathwani, 1989; Horvitz, Breese, Heckerman, Hovel, & Rommelse, 1998; Jameson, 1996; Vegas, 1995). Several studies conducted on the comparison of machine learning techniques used in student modeling systems show that Bayesian networks are effective for meeting the student modeling challenge (Gonzalez, Burguillo, & Llamas, 2006; Hamalainen & Vinni, 2006; Kotsiantis, Pierrakeas, & Pintelas, 2003; Minaeli-Bigdoli, Kashy, Kortemeyer, & Punch, 2003).

Hence, a dynamic Bayesian inference mechanism for students modeling has been proposed in this study. This mechanism is based on combined reasoning, and the student models are based on a data-driven approach (Graf, 2007).

**Student Modeling System:** The proposed system records student actions on the interactive learning content and feeds the data into the student modeling system for dynamic evaluation and feedback. The records kept about students are suitable for labeling. The proposed Bayesian student modeling system first simulated virtual students interacting with virtual learning content, and the simulation results were evaluated in advance (Aslan & İnceoğlu, 2008). Virtual students and virtual learning scenarios are frequently used in e-learning systems to test proposed solutions because access to actual students is not always convenient (Conati, Gertner, & Vanlehn, 2002; Liu, 2006; Mertz, 1997; Vanlehn & Niu, 2001). The conditional probability table values of nodes in the Bayesian network were assigned manually in parallel to results achieved in the simulation process. The student modeling system creates student profiles with respect to the Index of Learning Styles Questionnaire (Felder & Solomon, 1991). Only two of the four dimensions of the FLSM were considered during the research. It would be impractical to develop course materials reflecting every combination of learning style dimension because the number of learning scenes grows exponentially as greater numbers of learning dimensions are included (Felder, 1993). The knowledge input (visual/verbal) and knowledge processing (active/reflective) dimensions were used in this study.

The Bayesian student modeling layer used in the research procedure collects data from three main evidence nodes. These are: the learning scene itself, the usage of mouse and keyboard, and the period of time that the learning scene is active.

Incoming data from evidence nodes are processed by a combined inference method in which both predictive reasoning and diagnostic reasoning are used simultaneously. The conditional probability values of evidence nodes are assigned in parallel to the simulation results.

The most important problem after assigning values to nodes is the labeling of incoming data from evidence nodes. Labeling is required for mouse–keyboard use and the learning screen playtimes. Different threshold values are required for both nodes, with labels like “more–less” and “short–long.” The system was run with only four students in the startup phase for initialization of threshold values. After determination of the threshold values of evidence nodes, the system was run with students of the experimental and control groups. In addition to the initialization stage, the threshold values were dynamically updated with every student processed in the system.

**Browser-Shell Software:** A basic software solution has been developed to display learning scenes to students and to facilitate interaction. The system records the timer feedbacks incoming from mouse, keyboard, and screen activity recorders so that the data can be fed into the Bayesian updater. The browser-shell software separately records in a database all the data for every learning scene.

**Course Content:** Equivalent course content was designed for the experimental and control groups to test this study’s hypothesis. The control group follows a 27-scene learning content, while the experimental group follows a 46-scene content organized from a total of 98-scene learning-styles-based learning scenes.
Questionnaire Adaptation: Four main studies in the literature regarding the verification and validation of the FSLSM (Felder & Spurlin, 2005; Litzinger et al., 2005; Van Zwanenberg et al., 2000; Zywno, 2003) are based on the ILS questionnaire. Although several studies have addressed the English version of the ILS, there is only one study on the Turkish version (Samancı & Keskin, 2006). It has been emphasized that the questionnaires used in research studies should be designed in the native language of students to achieve more precise validation and verification analysis results (Hambleton & Patsula, 1999). Considering the scores on the Turkish-adapted version of the ILS, we decided that a re-adaptation of the ILS to Turkish could give rise to improvement on scores. All 44 items of Felder and Solomon's ILS were translated into Turkish and adapted accordingly (Atman et al., 2009).

Academic Achievement Test: A 24-item, multiple-choice academic achievement test was prepared to compare academic success levels between the control and experimental groups after tutoring sessions. This test was prepared by an academician who is an expert in teaching English as a foreign language, and in curriculum and instruction. The test was applied to both groups as a pre- and post-test.

Model
A quasi-experimental pre-test, post-test design was used in this study. The quasi-experimental design approach is frequently used as a mandatory choice in learning environments because of the challenges in applications with real students (Charles & Mertler, 2002; Creswell, 2002; McMillan, 2004; von Davier, Holland, & Thayer, 2004).

Research Group
The study was conducted with 46 undergraduate and graduate student volunteers from different departments of Ege University and İzmir Institute of Technology, with 23 assigned to the control group and 23 to the experimental group.

Procedures
The experimental design of this study consists of a control group and an experimental group. The learning process in both groups was conducted through individual appointments. The groups were made equivalent in academic achievement before the experimental procedure's application. The experimental group was tutored with an adaptive learning system based on learning styles, while the control group was tutored with a conventional computer-aided learning approach. Experimental procedures were realized over a period of 8 weeks between May 15th and July 14th of 2009. Appointments with the participants were arranged for an eight-week treatment, and an individual learning setting was provided. All the students in the control group were tutored with the same learning content, while the students in the experimental group were tutored in approximately 30-minute sessions with adaptive learning content based on their dynamically updated profiles.

Results
The data were evaluated in two main sections. In learning-style diagnosis performance, to what extent the system was able to diagnose the style of the learner was investigated. The difference between pre- and post-test scores was also studied to evaluate the participants' academic achievement.

Learning Styles Diagnosis Performance
Although several approaches are based on the FSLSM and Bayesian student modeling system, criteria related to learning style diagnosis performance have not been established. The criteria suggested by Garcia, Amandi, Schiaffino, and Campo (2007), which were accepted and used by several researchers, were also used in this study, and the results are compared with similar studies (Garcia et al., 2007; Garcia, Schiaffino, & Amandi, 2008; Graf & Kinshuk, 2008; Özpolat & Akar, 2009). The diagnosis performance was calculated as 41% in the visual/verbal dimension and 54% in the active/reflective dimension.

Academic Achievement
The experimental and control groups had comparable performances in the pre-test. The analysis of within-group comparison revealed that both the experimental and control groups achieved higher scores in the post-test. When the means for the pre- and post-test were compared, there was a significant difference for both groups toward the positive. However, no significant difference was found between the control and experimental groups after the treatment.
Discussion

A computer assisted learning middleware was developed to evaluate a machine-learning based student modeling system in this study. The ILS, which is the scale of the FSLSM model, was adapted into Turkish. A grammar topic in teaching English was transferred into an e-learning setting, and an English achievement test was used in the study.

The study may have contributed to the literature as it targeted language learning using a Bayesian-based student modeling system. The use of mouse and keyboard and time spent using them might also have contributed to the investigation of evidence nodes’ effectiveness. A solution was sought for the possible concept drift problem in a machine based user, also discussed by Webb et al. (2001). A dynamic student modeling system that can update itself against concept drift and adapt itself according to the changes in learning styles during the learning process was suggested.

The ILS scale was translated into the native language of the study participants. The reliability and validity of the scale increased in this Turkish version compared to the Turkish version used in another study.

Because learning settings arranged according to learning styles increase academic achievement, a student modeling system based on learning styles in an e-learning setting was developed. The performance of machine learning based student modeling system to diagnose learning styles was evaluated. According to the style diagnosis performance criteria (Garcia et al., 2007), the performance level was 41% for the active/reflective dimension and 54% for the visual-verbal dimension. This performance result was compared to similar studies in the literature.

The student modeling system was able to diagnose 50% of the verbal learners and 75% of the active learners. There has not been any correct diagnosis for verbal and active learners in previous studies. Compared to other studies, the findings of this study seem promising. In addition, there has not been an experimental study that could diagnose at least one student correctly across all learning styles mostly because the performance analysis criterion developed by Garcia et al. (2007) assigned half points for medium-level estimations. In previous similar studies conducted with engineering students, the suggested systems performed better for active- and visual-learner styles. However, the performance level might decrease in estimation when applied to students from different departments. In this study, the context was formed for English language learning, which is an interdisciplinary domain, and thus, students from different departments shared the same educational setting. Therefore, we must emphasize that studies with broader participation, including different learning styles, and data collected from more evidence nodes are needed to discuss how successful the student modeling systems are.

In addition, a pre-test, post-test quasi-experimental design was used to investigate whether the student modeling system based on a computer-assisted learning setting affected academic achievement compared to a conventional computer-assisted learning setting. For the within-group analysis, both the experimental and control groups achieved higher scores on the post-test. The difference between pre- and post-test scores was statistically significant. When the two groups’ pre- and post-test results were compared, there was a non-significant difference. As a result of the findings, the hypothesis “Learning styles based tutoring results in an increase in academic achievement with respect to traditional web-based tutoring” was rejected.

The ILS scale was translated into the native language of the study participants. The reliability and validity of the scale increased in this Turkish version compared to the Turkish version used in another study.

Because learning settings arranged according to learning styles increase academic achievement, a student modeling system based on learning styles in an e-learning setting was developed. The performance of machine learning based student modeling system to diagnose learning styles was evaluated. According to the style diagnosis performance criteria (Garcia et al., 2007), the performance level was 41% for the active/reflective dimension and 54% for the visual-verbal dimension. This performance result was compared to similar studies in the literature.

The student modeling system was able to diagnose 50% of the verbal learners and 75% of the active learners. There has not been any correct diagnosis for verbal and active learners in previous studies. Compared to other studies, the findings of this study seem promising. In addition, there has not been an experimental study that could diagnose at least one student correctly across all learning styles mostly because the performance analysis criterion developed by Garcia et al. (2007) assigned half points for medium-level estimations. In previous similar studies conducted with engineering students, the suggested systems performed better for active- and visual-learner styles. However, the performance level might decrease in estimation when applied to students from different departments. In this study, the context was formed for English language learning, which is an interdisciplinary domain, and thus, students from different departments shared the same educational setting. Therefore, we must emphasize that studies with broader participation, including different learning styles, and data collected from more evidence nodes are needed to discuss how successful the student modeling systems are.

In addition, a pre-test, post-test quasi-experimental design was used to investigate whether the student modeling system based on a computer-assisted learning setting affected academic achievement compared to a conventional computer-assisted learning setting. For the within-group analysis, both the experimental and control groups achieved higher scores on the post-test. The difference between pre- and post-test scores was statistically significant. When the two groups’ pre- and post-test results were compared, there was a non-significant difference. As a result of the findings, the hypothesis “Learning styles based tutoring results in an increase in academic achievement with respect to traditional web-based tutoring” was rejected.

The reason for the non-significant statistical result in achievement between the two groups may have resulted from disadvantages of the pre-test, post-test quasi-experimental study design, which lacks the element of random assignment to treatment. As volunteer participants are motivated and willing individuals, they may not reflect the population in a real sense (Best & Kahn, 2006). Quasi-experimental design and volunteer participation may have affected this study's findings.

Much research has been conducted on the achievement of learning based on learning styles, as compared to traditional learning. However, the effects on achievement of a student modeling system based on learning styles in computer-assisted settings has not been studied. In this study, an experimental design compared computer-assisted learning settings wherein student modeling was used and not used. The interdisciplinary context of the study may also contribute to related literature. For further studies, e-learning systems that use student modeling systems can be compared to others for effects on achievement.

Finally, we should remember that student modeling and e-learning systems were developed to help people (Frank, Reich, & Humphreys, 2003). The techniques in machine learning used in e-learning systems and student modeling systems should also be evaluated in terms of ethics (Aiken & Epstein, 2000), and author royalties (Kennedy, 2002) should also be considered.
References/Kaynakça


