KURAM VE UYGULAMADA EĞİTİM BİLİMLERİ EDUCATIONAL SCIENCES: THEORY & PRACTICE

Received: January 24, 2016
Revision received: November 24, 2016
Accepted: September 14, 2017
OnlineFirst: November 10, 2017

www.estp.com.tr
DOI 10.12738/estp.2017.6.0059 • December 2017 • 17(6) • 1893–1910

Research Article

Learning Analytics of Student Participation and Achievement in Online Distance Education: A Structural Equation Modeling*

Mustafa Koç[1]
Süleyman Demirel University

Abstract

This study proposed a theoretical model explaining causal relationships between student participation and academic achievement through their learning analytics in a web–based distance education course, testing it using structural equation modeling (SEM) with an empirical dataset. The sample was composed of 167 university students enrolled in the Department of Computer Programming in the Distance Education Vocational School at a major state university in Turkey. Student participation was operationalized through learning analytics of the number of submissions to discussion forums and attendance to online lectures whereas academic achievement was represented by students' performance on their project assignments and final exam. The results of the SEM analysis indicated that the modified version of the model had a good fit with the data (x^2=2.52, *df*=1, *p*>.05, x^2/df=2.52, GFI=.99, CFI=.99, RMSEA=.09, SRMR=.03). Discussion forum submission and online lecture attendance were found to be positively associated with each other. They had a positive direct effect on students' project scores and a positive indirect effect on students' final exam scores via their project scores. Moreover, discussion forum submission was found to have a direct positive effect on students' final exam scores. Practical implications and suggestions for further research are discussed within the context of online learning.

Keywords

Online distance education • Learning analytics • Student participation • Academic achievement • Structural equation modeling

* A part of this study was presented as an oral presentation at the 8th International Computer Education and Instructional Technologies Symposium, September 18-20, 2014, Trakya University, Edirne, Turkey.

1 **Correspondence to:** Mustafa Koç, Department of Computer Education and Instructional Technology, Faculty of Education, Süleyman Demirel University, Isparta 32260 Turkey. Email: mustafakoc@sdu.edu.tr

Citation: Koç, M. (2017). Learning analytics of student participation and achievement in online distance education: A structural equation modeling. *Educational Sciences: Theory & Practice, 17,* 1893–1910. http://dx.doi.org/10.12738/estp.2017.6.0059

Truly remarkable is the ever increasing growth and impact of information and communication technologies (ICT) on human life. Through e–commerce, e–government, e–entertainment and the like, ICT innovations have dramatically changed the way the daily activities and important tasks are undertaken, and education has not been destitute of such changes. Scholars and research groups have been investigating emerging educational technologies and their possible impacts on teaching and learning. Recent influential reviews and projects such as New Media Consortium's Horizon Reports, A Roadmap for Education Technology funded by the National Science Foundation in the USA, and the Technology Outlook for UK Tertiary Education 2011–2016 Report describe emerging technologies that are to gain dominance and significance in education along with key trends and critical challenges (Ng'ambi, 2013; Spector, 2013). Continually reported in these reports as one of the educational technologies that are most likely to affect teaching and learning in the short and medium term (one to five years) is learning analytics. In fact, learning analytics has been lauded in New Media Consortium's latest report (Johnson et al., 2016) as one of the most important emerging trends thought to accelerate technology adoption in higher education in a one–year or less time horizon. Its usage has the potential to help educational institutions in student retention, easing the burden of accountability, providing personalized learning experiences, and increasing student success (Dietz–Uhler & Hurn, 2013). This study therefore focuses on the use of learning analytics components that can be obtained from web–based distance education environments and on investigating its potential for improvement in student learning.

Learning Analytics

Although the field of learning analytics is still in its early stages, two research communities, i.e., The Society for Learning Analytics Research (SOLAR) and International Educational Data Mining Society (IEDMS), have already established with their own academic conferences and journals. These organizations collaboratively work to share common definitions, research, methods, and tools for data mining and analytics (Baker & Siemens, 2014). SOLAR defines learning analytics as the measurement, collection, analysis, and interpretation of data related to learners' behaviors and learning contexts in order to optimize instructional processes and environments (Siemens & Gasevic, 2012). It occupies the intersection of educational research and computational techniques to capture and analyze learners' data (Fırat & Yüzer, 2016; Knight, Buckingham Shum, & Littleton, 2014). The field connects to and builds on several disciplines, including but not limited to educational sciences, instructional design and technology, computer science, user modeling, advanced statistics, and information visualization (Demirbaş & Koç, 2015; Gasevic, Dawson, Mirriahi, & Long, 2015). Although the field of education has traditionally dealt with data and analysis, the integration of ICT into education makes it now possible to record, retrieve, and store aggregate and large sets of digital data (i.e., educational data mining) in an easy and cost–effective manner. The information retrieval technologies also allow for gathering incidental, unstructured, complex, and various data that is highly authentic in terms of user behavior. Learning analytics as an emerging research domain explores the systematic use of such data to ultimately refine pedagogical strategies, regulate institutional costs, identify and help struggling students, assess factors affecting student success and teacher efficacy, and improve the assessment of student performance (Larusson & White, 2014). Since valuable educational data have become available for analysis and interpretation, institutions have started implementing data warehouses for not only improving student learning but also organizational decision making and preparation for a future in which analytics will have become a strategic asset (Stiles, Jones, & Paradkar, 2011).

Given that learning analytics is still in its infancy, the literature germane to its process is quite scarce. Comparing and combining several representations of analytical process in various disciplines, Elias (2011) put together an ongoing model of learning analytics with three–phase cycles: data gathering, information processing, and knowledge application. Typically, the process starts by capturing raw data about learning as a social activity. This data might consist of a variety of information, ranging from learners' demographic characteristics to patterns of interaction with peers and course content. Afterward, selected data are analyzed using advanced statistical methods (e.g., regression analysis, path analysis etc.) and visualization techniques. The ultimate aim here is to develop predictive models of optimum learning conditions. As a result of this analysis, the final phase comprises using and refining effective learning activities. In a similar vein, Pardo (2014) synthesized generic techniques and algorithms previously used in other research domains, offering a five–stage process for the deployment of learning analytics experience. The first stage, *capture*, refers to the collection and storage of student data. The second stage, *report*, corresponds to the summarization and visualization of the collected data through arbitrarily complex processing methods. The outcomes are delivered to stakeholders as well. The third stage, *prediction*, involves the utilization of prediction techniques to answer frequently encountered educational problems (e.g., a probability of a student failing a course). The fourth stage, *act*, includes the design, development, and implementation of automated or manual solutions based on the consequences of prior prediction. The final stage, *refinement*, assesses the efficacy of resulting actions and modifies them so that they may be viable in the long–term. Adopting a holistic account, Pardo (2014) presented each stage as being bounded up with earlier stages rather than being a single step in the process.

One implication of learning analytics lies in the development of personalized learning context and recommender systems. As Spector (2013) exemplifies, a customized learning activity can be suggested to a particular student with an instructional struggle based on the analysis of interests, preferences, actions, and previous performance of similar students who have experienced the same struggle but who succeeded with the help of a supporting instructional activity. Revealing the relationships among personal and contextual data can facilitate the configuration of particular learning experiences that are not only relevant but also interesting to students with different profiles. This notion brings a paradigmatic shift from institutional perspectives to learner and teacher concerns in educational settings, as it is no longer educational authorities and funders who benefit most from analytics, but rather students and teachers (Long & Siemens, 2011). Moreover, learning analytics components allow for more reliable formative assessment and program evaluation as they use students' actual behaviors instead of their self–reports gathered through questionnaires (Hung, Hsu, & Rice, 2012). These evaluation outcomes support reflection and prediction (Greller & Drachsler, 2012). Teachers and students can reflect upon their practices based either on their own self–knowledge or on each other's datasets. Predictive models can be developed to identify the values of learning activities for explaining academic failure and success. In accordance with the increasing use of ICT in education, it is expected that learning analytics will play a critical role in collecting evidence of improvement and in building awareness of progress in developing 21st century skills (Kong et al., 2014). Therefore, a number of major universities have already established special committees, allocating the necessary budget to explore what learning analytics holds for education in the near future (Yıldız & Bahçeci, 2014). They have been using learning analytics tools to provide dashboards for both students and instructors in order to track learning progress in their courses (Dietz–Uhler & Hurn, 2013).

The proliferation of online learning environments, such as games, virtual worlds, simulations, massive open online courses (MOOCs), and learning management systems (LMSs) (e.g., Moodle and Blackboard), in educational environments, especially in web–based distance education contexts, has facilitated the use of learning analytics (Fırat & Yüzer, 2016). LMSs enable learners to access digital course resources, complete learning tasks, and interact with others while also allowing instructors to check learners' actions and performance. Their most important contribution to learning analytics is that they offer the ability to create web logs that include various trails of student behavior (i.e., digital footprints). These log files record a large amount of data related to learners' profiles, social networks, learning tendencies, sharing and updates, usage of course modules (e.g., assignments, forums, and chat), performance indicators, and so on. There are also special software tools (e.g. SNAPP and LeAR) that work concordantly with some LMSs to analyze and visualize such data (Yıldız & Bahçeci, 2014). Consequently, the availability of large data on the latest digital technologies has changed the way that evidence is gathered and interpreted in education. Educators are no longer limited and burdened by the development and implementation of measurement tools. The relevant software collects the necessary data inherit in student learning and then analyzes and presents them in the forms of dashboards and graphical visualizations (Cope & Kalantzis, 2015). All these tasks are completed through technology mediation without user intervention.

Since the field is still in its early stages of research and practice, the literature is not replete with studies on the application of learning analytics. Another reason for the limited number of existing studies might be the complex and interdisciplinary nature of studying learning analytics, as it requires not only dealing with massive data and collaborating with an extensive team of experts but also that various methods and skills (Fırat & Yüzer, 2016). Koç (2016) examined current research trends and issues on learning analytics by reviewing 41 empirical articles published in refereed leading journals from 2010 to 2015. He found that the number of studies began to increase rapidly as years passed and that the majority of them focused on computer, science, and engineering subjects conducted in higher education institutions where distance education and e–learning tools are widely used. The number of articles published each year was double the number published the preceded year. Koç (2016) revealed that while a little more than half of the studies aimed to explore and model the relationships between learning analytics and learning outcomes, the remaining aimed to investigate the role of learning analytics in the management and evaluation of learning processes. The former group of studies focused on prediction whereas the latter group focused on monitoring. He further discovered that the majority of the data used in the studies included learner behaviors and gains that were automatically saved in the log files and were analyzed through such advanced techniques as structural equation modeling, cluster analysis, social network analysis, algorithms, etc. Koç (2016) also found that almost all studies were carried out by multiple researchers.

Joksimovic, Gasevic, Kovanovic, Riecke, and Hatala (2015) showed that certain indicators of social presence (e.g., frequency of logging into discussion forum, number of posts) were significant predictors of final grades in a master's level online computer science course. While Jo, Kim, and Yoon (2015) treated total login time, login frequency, and regularity of login interval as indicators of adult learners' time management strategies in an LMS, they identified (ir)regularity of the login interval as the significant predictor of learning performance. In an introductory economy course, computer–assisted formative assessments (tracking data from e–tutorial systems) were found to be the best predictor for detecting underperforming students and academic performance whereas basic LMS data did not substantially predict learning (Tempelaar, Rienties, & Giesbers, 2015). A couple of studies demonstrated the prediction or detection of dropout rate/risk levels in open and distance education programs from some students' LMS usage, demographic, and course registration data (Cambruzzi, Rigo, & Barbosa, 2015; Grau–Valldosera & Minguillon, 2014; Yasmin, 2013). In a location–based mobile learning game, providing students with visualization of their gaming behaviors (e.g., time spent to reach geographical zones, levels passed, scores obtained) were found to be beneficial for the diagnosis of their own performance (Melero, Hernandez–Leo, Sun, Santos, & Blat, 2015).

As far as those studies investigating the role of learning analytics in monitoring student learning progresses are concerned, implementing relevant learning analytics in teaching and learning was shown to foster self–regulated learning and improve time management skills (Tabuenca, Kalz, Drachsler, & Specht, 2015), enhance the evaluation of reading literacy skills (Picher & Ebner, 2015), support learner profiling for the improvement and individualization of the learning environment (Taraghi, Saranti, Ebner, Müller, & Grobmann, 2015), lead to better teacher ability to diagnose problems concerning participation of students (van Leeuwen, Janssen, Erkens, & Brekelmans, 2014), provide awareness about the emotions of the learners to the instructors (Leony, Munoz–Merino, Pardo, & Kloos, 2013), obtain insight into the learning process, and create methods to enhance or evaluate learning occurred in virtual learning environments, such as online laboratories (Qvist et al., 2015).

The Present Study

This study aims to build a theoretical model that explains the causal relationships between student active participation and academic achievement through their learning analytics in a web–based distance education course and then test it with an empirical dataset. Investigating the benefits of students' active participation for academic success in online courses has become a popular research trend in the related literature (Rocca, 2010). For the hypothesized model (Figure 2) in the present study, student participation was operationalized through learning analytics of the number of submissions to discussion forums and attendance to online lectures. Academic achievement was represented by the learning analytics of project assignments and final exam scores. The theoretical foundation of the model relies on the community of inquiry (CoI) framework and social constructivist learning.

The CoI framework has been shown to be a useful approach to understand and develop online learning environments and strong predictor of learning outcomes (Arbaugh, 2008). It focuses on creating a community of learners with three overlapping dimensions: social presence, teaching presence, and cognitive presence (Garrison, Anderson, & Archer, 2000). Meaningful learning occurs at the intersection of these dimensions through social interactions and educational experiences. Regarding online education, social presence entails immediacy, connectedness, group cohesion, open communication, collaboration, help, and sharing. Using e–mail, forums, blogs, instant messaging, and video conferencing technologies, learners can enhance social presence as they communicate with their peers, teachers, and other professionals. Online learners with a stronger sense of community and greater cognitive learning in a positive socio–emotional climate are expected to feel less isolated and more satisfied with academic programs (Rovai, 2002). Expectedly, prior research reveals that social presence is significantly associated with student learning and satisfaction (Garrison & Arbaugh, 2007; Shea, 2006; Williams, Duray, & Reddy, 2006). Social encounter is also seen as the driving force behind understanding and meaning making in social constructivism. Through his familiar notion of zone of proximal development (ZPD), Vygotsky (1978) states that learners, with the assistance of more knowledgeable ones, can perform more actions and develop more profound comprehension than what they can do on their own. Therefore, online dialogue and discussion can mediate ZPD by providing learners with feedback, support, and negotiation on learning tasks.

Teaching presence reflects the roles of course instructor with regards to design, facilitation, and direction of cognitive and social activities (Garrison et al., 2000). Having learners interact with others and course content is not enough for meaningful learning. Instructors should define and structure such interactions. They can accomplish this by providing guidelines on how to make submissions to discussion forums, developing audio/video lectures, preparing online course content, moderating student discussion, using various means of feedback, and so on (Garrison & Arbaugh, 2007). Consistent with social constructivist learning, these teacher interventions contribute to scaffolding learners within their ZPD and thus enabling them to perform better. There is also research evidence indicating the significant role played by teaching presence in successful online learning (Arbaugh, 2008; Shea, Li, & Pickett, 2006). Since the instructor can intervene directly in the course's content and social interactions, teaching presence also aids in shaping both social and cognitive presence based on the types of instructional media and environments (Kupczynski, Ice, Wiesenmayer, & McCluskey, 2010).

Cognitive presence is grounded in learners' engagement in course content and can be developed through four phases: (i) identification of an issue or a problem for further inquiry, (ii) learners' exploration of this issue, (iii) learners' construction of meaning from the ideas developed during exploration, and (iv) learners' application of newly gained knowledge (Garrison et al., 2000). It is complementarily related to teaching presence because of the instructor's crucial role in designing and structuring course content and learner interaction (Arbaugh, 2008). The analysis of online discussion, performance on learning tasks (e.g., assignments, projects), and grades can be used to reflect cognitive presence in online learning (Akyol & Garrison, 2011). Previous studies suggest a strong relationship between cognitive presence and learning outcomes (Akyol & Garrison, 2011; Arbaugh, 2008).

Based on the aforementioned relationships among the three presences within the context of online learning, the model presented in this study (Figure 2) proposes the following research hypotheses:

1. Submission to discussion forums is positively associated with online lecture attendance.
2. Submission to discussion forums has a direct positive effect on the performance in project assignments.
3. Attendance to online lectures has a direct positive effect on the performance in project assignments.
4. Performance on project assignments has a direct positive effect on final exam performance.
5. Discussion forum submission has an indirect positive effect on final exam performance through performance on project assignments.
6. Attendance to online lectures has an indirect positive effect on final exam performance through performance on project assignments.

Method

Participants

The participants were made up of 167 university students enrolled in a two–year undergraduate program in the Department of Computer Programming in the Distance Education Vocational School at a major state university in Turkey. Students were not selected by means of any sampling procedure, but include all those (i.e., entire population) enrolled in the program at the time of the study. Their program of study was fully based on distance education and carried out through a web–based LMS. Therefore, they were not from a specific location but rather from all over the country. Of the sample, 106 (%63) were male and 61 (%37) were female distance education students. The sample size (n=167) was considered sufficient for this study because it met Kline's (2005) recommended value of 100–150 cases to obtain reliable results in structural equation modeling studies.

Research Procedures

The study was designed as a descriptive correlational research as it explores the multiple relationships among student participation and learning outcome indicators in a web–based distance education course (Büyüköztürk, Çakmak, Akgün, Karadeniz, & Demirel, 2011). It took place in the General and Technical Communication course that was taught by the researcher. This online course was managed via a commercial LMS. On this online platform, students were provided with course contents, discussion forums, project assignments, and weekly online lectures. The researcher gave these synchronous lectures using Adobe Connect's web conferencing software in a specially–designed studio at the university (Figure 1). This software was embedded in the LMS so that students would be able to attend to web conferencing sessions from their current locations via the Internet. It enabled the researcher to share his video, audio, and course content (e.g., PowerPoint presentations, lecture notes etc.) and to allow students to share their own videos, audios, and content. There was also a text–based instant messaging screen through which students could chat with everyone in the course.

Figure 1. Video Conferencing Studio (on the left) and Software (on the right).

The hypothesized model was constructed as a recursive structural model and tested using structural equation modeling (SEM) path analysis in LISREL 8.80 software (Çokluk, Şekercioğlu, & Büyüköztürk, 2010). Descriptive statistics and correlation coefficients were computed using SPSS 13.0 software. Submission to discussion forums and attendance to online lectures were treated as exogenous variables whereas performances on both project assignments and the final exam were treated as endogenous variables. The social, teaching, and cognitive presences explained above were developed by operationalizing these variables.

Measures

Discussion forum submission. Using the asynchronous discussion forum of the LMS, the researcher opened discussion topics related to the definitions, examples, and barriers of various communication types and asked students to make comments. Therefore, this variable was measured by the number of total posts that each student sent to discussion forums and was employed as a learning analytics for student participation.

Online lecture attendance. During the 14–week teaching period, online lectures were given on a weekly basis through video conferencing. Each lecture lasted approximately 75 minutes. Students were encouraged, though not mandated, to attend the lectures. Students' attendance was automatically recorded in the LMS's web logs. Thus, the variable was operationalized by determining the number of weeks each student participated in online meetings. It was employed as another learning analytics for student participation.

Project score. As a part of the course assessment, students were required to complete two main project assignments and submit them to the LMS. The tasks in the first and second projects required student to develop a business letter and visual communication tools (e.g., form, table, graph, and diagram), respectively. Projects could be done on any subject matter of interest to students. The researcher submitted descriptions and requirements of the tasks on the LMS and provided students with guidance when they needed. Each project was scored on a 50–point Likert–type rubric. The researcher adapted this rubric from a variety of generic rubrics used in project assessments and focused on such characteristics as originality, suitability, organization, effective use of communication elements, and both timely and correct submission. Hence, a composite variable ranging from 0 to 100 was created by summing up the scores of the two projects for each student and was treated as a learning analytics for student achievement.

Final exam score. In order to make the course assessment as reliable and valid as possible, students were invited to the university campus during the last weekends of the semester to take a written and face–to–face final exam. The score taken from this exam had composed the greatest part of the overall course grade. The exam consisted of 30 multiple–choice questions addressing important topics covered in the online course. All questions were developed by the researcher and subjected to expert review to assure both content and face validity. The Kuder–Richardson (KR–20) internal consistency coefficient for all questions was .77, which suggests sufficient reliability. Students earned equal scores (3.33) for each correct answer. Consequently, the final exam score was measured as a composite variable by summing up the earned scores from all questions. It was treated as another learning analytics for student achievement.

Results

Descriptive Findings and Inter–relationships

Table 1 shows minimum and maximum values, means, standard deviations, skewness and kurtosis values, and Pearson's simple correlation coefficients with regards to study variables. As can be seen, discussion forum submission is significantly and positively associated with online lecture attendance (r=.30, p<.01), project score (r=.50, p<.01) and final exam score (r=.37, p<.01). Online lecture attendance is significantly and positively associated with project score (r=.48, p<.01). Project score is significantly and positively associated with final exam score (r=.27, p<.01). However, there was no significant association between online lecture attendance and final exam score.

Table 1

Descriptive Statistics and Correlations among the Variables

Variable	Range	M	SD	Skewness	Kurtosis	1	2	3	4
1. Discussion forum submission	0–3	.78	1.02	1.01	–.30	1.00	.30*	.50*	.37*
2. Online lecture attendance	0–14	3.12	3.12	1.10	.15		1.00	.48*	.07
3. Project score	0–100	50.30	45.56	–.11	–1.80			1.00	.27*
4. Final exam score	26.6–99.9	64.95	13.05	–.23	.02				1.00

*p < .01

Findings of Structural Equation Modeling

Before testing the hypothesized model, the assumptions of SEM analysis were examined. The data set did not have any missing values. The skewness and kurtosis values in Table 1 as well as visual inspection of normal probability plots indicated that variables could be accepted as normally distributed. The examination of boxplots indicated a few suspected univariate outliers. However, they were not disruptive to distributions of the variables because the means were not very different from the 5% trimmed means. Mahalanobis distance values were calculated for the inspection of multivariate outliers (Tabachnick & Fidell, 2007). There were no multivariate outliers since all cases had Mahalanobis values lower than the critical Chi–square value of 16.27 (df=3, *alpha*=.001). Multicollinearity did not occur in the data because inter–correlations among the variables (Table 1) were below the threshold value of .85 (Kline, 2005). The sample size required for SEM analysis is a controversial issue. The related literature suggests considering several issues when deciding sample size, which are model specification (including relevant variables in the model), model size (complexity of the model), normality of the data, and estimation method used (Teo, 2010). A bigger sample size should be used with more complex models that include more variables and with those datasets violating the normality assumption. Moreover, several researchers recommend that the minimum sample size to use maximum likelihood estimation procedure be between 100 and 150 participants (Ding, Velicer, & Harlow, 1995; Kline, 2005). Hence, the sample of this study (n=167) is acceptable since the hypothesized model is quite straightforward and its dataset is normally distributed.

Because the data met the assumptions, a SEM analysis was conducted using maximum likelihood estimation. The hypothesized model did not have acceptable fit to data (χ^2=12.69, df=2, p<.01, χ^2/df=6.35, GFI=.96, CFI=.91, RMSEA=.18, SRMR=.07). Figure 2 presents the path coefficients and explained variances in the model.

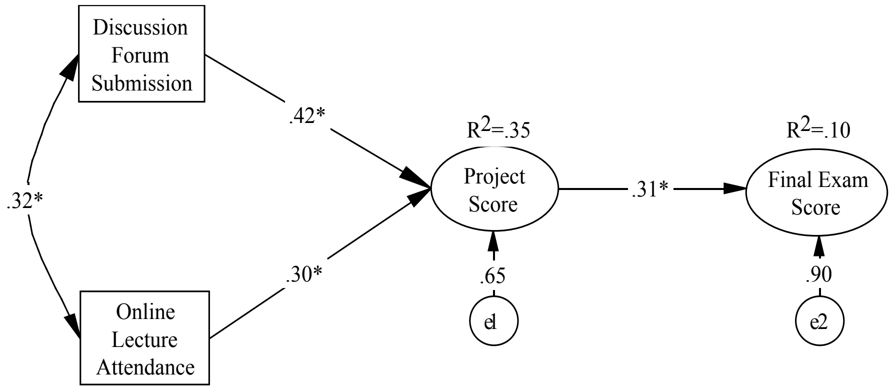

Figure 2. Path Analysis of Hypothesized Model (*p<.01)

The modification indices in the LISREL output recommended that adding a path from discussion forum submission to final exam score would result in significant amount of decrease (10.3) in the Chi–square value. Since this recommendation was justifiable with the theoretical framework of the study, the hypothesized model was modified by adding this path.

Again, the modified model was tested through a SEM analysis using maximum likelihood estimate. The results indicated a good fit this time (χ^2=2.52, df=1, p>.05, χ^2/df=2.52, GFI=.99, CFI=.99, RMSEA=.09, SRMR=.03). In fact, the model comparison indices for the modified model (AIC=20.52, CAIC=57.58, ECVI=.13) were smaller than those for the initial hypothesized model (AIC=28.69, CAIC=61.63, ECVI=.17), suggesting a better fit of the modified model to data (Hu & Bentler, 1995).

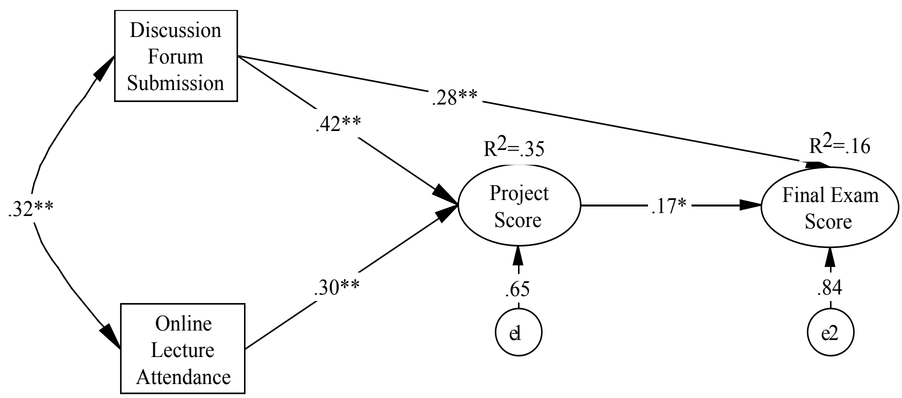

Figure 3. Path Analysis of Modified Model (*p<.05, **p<.01)

The standardized path coefficients in the modified model (Figure 3) clearly show that discussion forum submission and online lecture attendance were positively associated (β=.32, p<.01) and that both had positive direct effects on project score (β=.42, p<.01 and β=.30, p<.01, respectively), explaining 35% of its variance. Final exam score was positively and directly affected by project score (β=.17, p<.05) and discussion forum submission (β=.28, p<.01). The indirect effects of exogenous variables were computed by multiplying related path coefficients (Çokluk et al., 2010). As hypothesized, final exam score was positively and indirectly affected by discussion forum submission (β=.07, p<.01) and online lecture attendance (β=.05, p<.01) via project score. Consequently, 16% of the variance in final exam scores was explained by discussion forum submission, online lecture attendance, and project score.

Discussion and Conclusion

With theoretical guidance of the CoI framework and social constructivist learning, the purpose of this study was to develop a structural model that tests and estimates casual relationships between student participation and learning achievement in an online course. The numbers of submissions on the discussion forum and online lecture attendance were used as learning analytics of student participation whereas project and final exam scores were employed as learning analytics of learning achievement. Turning back to the beginning of the study, the results supported all hypotheses on the direct effects that discussion forum submission and online lecture attendance had on the project score as well as the indirect effects that they had on the final exam score. Additionally, the modified model showed that discussion forum submission had a direct effect on one's final exam score. Overall, all these findings support the determining role of social, teaching, and cognitive presences in student achievement in online learning environments (Arbaugh, 2008; Garrison et al., 2000; Garrison & Arbaugh, 2007). They are also consistent with previous studies indicating that greater participation in online discussions and interaction may lead to higher grades (Handelsman, Briggs, Sullivan, & Towler, 2005; Tayebinik & Puteh, 2013).

Among the variables representing student participation in this study, discussion forum submission was found to be the strongest predictor for learning outcomes due to both its direct and indirect effects with higher standardized regression weights on the project and final exam scores. One possible explanation for this can be that discussion boards allow for enhanced social context and course involvement including group discussion, student–student interaction, and student–instructor interaction. Although online live lectures can provide the same social context, there is an overall tendency to utilize them to transfer information in a one–way communication, occurring most often from course instructor to students. Although this may increase cognitive presence to some extent, it may not promote social and teaching presence. Some underlying reasons for such utilization might include time limitations, a low level of student participation, and a lack of flexibility associated with synchronous online communication (Skylar, 2009). Text–based asynchronous communication, on the other hand, provides students with a more self–paced learning experience. They can access learning materials anywhere and anytime, spend more time in preparing their discussion posts, and externalize ideas through writing. Moreover, shy and less outspoken students can be more engaged in asynchronous modes of interaction. A few studies comparing asynchronous and synchronous interaction exist in the literature and suggest that both are equally effective in delivering online instruction (Skylar, 2009). However, there is insufficient empirical evidence for the comparison of their contributions to student learning within the same online course. The present study presents some preliminary findings in this regard and calls for further research to thoroughly explore the effects of different utilizations of asynchronous and synchronous interactions on learning outcomes.

As opposed to discussion forum submission, online lecture attendance has no significant direct effect on final exam scores. Although this was not hypothesized in the structural model, neither was it recommended in the modification indices. The lack of direct effect can be explained by the mediation effect of project score on the relationship between online lecture attendance and final exam score. It appears that if students with higher attendance to online lectures are more successful in the final exam, it is partly due to their high performance in the project assignments. This suggests that pedagogical approaches employed during online lectures should engage learners in active, contextual, and reflective learning. For example, students should be encouraged to go beyond being just passive listeners by means of problem–solving or project–based learning tasks in which they need to think, share, and use their knowledge. In this case, instructors should play the role of mentors and facilitators rather than information presenters.

Although both participation in the discussion forum and attendance to online video–conferencing lectures were found to contribute to students' achievement, both the mean and the range of the number of discussion posts and the mean number of lectures were very low. The most important reason behind these findings might be that students were not mandated for these tasks. Since the course through which this research was conducted was taught completely online, such an obligation might not be consonant with convenience and flexibility of distance education. It is well–known that students prefer online programs to accommodate their busy lives with multiple responsibilities including jobs, appointments, travels, family matters, and so on (Jaggars, 2014). Hence, synchronous meeting at a specific time may not be suitable for every student's working schedules, possibly leading to reductions in attendance. Students should be informed about the opportunities and advantages of joining video conferences (e.g., increasing the sense of instructor and peer support) so that they might be motivated to attend. Another reason for the low level of participation in the asynchronous discussion board could be that participation, or lack thereof, did not affect students' grade in the course. Furthermore, there might be students who are reticent or who find postings shallow. In order to encourage their participation, students can be given a rubric-based grade with brief guidelines and expectations for their participation on the forum.

On the whole, the results recommend that student success in online learning be promoted by increasing student participation in discussion forums and online lectures with more engaging learning activities. Future research should use the structural model in different conditions to confirm whether it pertain to a specific sample of students and courses. The results are limited to four learning analytics only. The model can be enhanced by integrating other student profiles and online learning behaviors. The measurement of student participation can also be improved by focusing more on knowledge tracking (e.g., quality of discussion posts) rather than activity tracking (e.g., quantity of discussion posts). The frequency and duration of the visits to the course's LMS can also be used as other measures. Another limitation of the study is the relatively small sample size, which works to reduce the generalization of the findings. Larger samples are preferred due to the massive data–driven aspect of learning analytics research. However, this may not be possible, especially at the beginning stage of learning analytics field, because of limited research conditions (e.g., low number of student enrollment, lack of student records, etc.). In fact, a recent meta analysis reported that half of the learning analytics studies published between 2010 and 2015 employed samples of 0 to 100 participants (Koc, 2016). Therefore, as the number of students increases and their learning records are recorded and stored, future studies should be conducted to test the model on a larger number of participants in order to increase external validity.

References

Akyol, Z., & Garrison, D. R. (2011). Understanding cognitive presence in an online and blended community of inquiry: Assessing outcomes and processes for deep approaches to learning. *British Journal of Educational Technology, 42*(2), 233–250.

Arbaugh, J. B. (2008). Does the community of inquiry framework predict outcomes in online MBA courses? *The International Review of Research in Open and Distance Learning, 9*(2).

Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/490/1045

Baker, R., & Siemens, G. (2014). Educational data mining and learning analytics. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 253–274). New York: Cambridge University Press.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2011). *Bilimsel araştırma yöntemleri* (10.Baskı) [Scientific research methods (10th ed.)]. Ankara: Pegem Akademi Yayıncılık.

Cambruzzi, W., Rigo, S. J., & Barbosa, J. L. V. (2015). Dropout prediction and reduction in distance education courses with learning analytics multitrail approach. *Journal of Universal Computer Science, 21*(1), 23–47.

Cope, B., & Kalantzis, M. (2015). Interpreting evidence–of–learning: Educational research in the era of big data. *Open Review of Educational Research, 2*(1), 218–239.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2010). *Sosyal bilimler için* çok *değişkenli istatistik: SPSS ve LISREL uygulamaları* [Multivariate statistics for social sciences: Applications of SPSS and LISREL]. Ankara: Pegem Akademi Yayıncılık.

Demirbaş, M. E., & Koç, M. (2015). Mobil teknoloji araçlarında öğrenme ve kullanıcı analitikleri: Bir alanyazın derlemesi [Learning and user analytics in mobile technology tools: A literature review]. *Proceedings of International Conference on Education in Mathematics, Science & Technology (ICEMST)*, 342–348.

Dietz–Uhler, B., & Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning, 12*(1), 17–26.

Ding, L., Velicer, W. F., & Harlow, L. L. (1995). Effects of estimation methods, number indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling, 2*, 119–144.

Elias, T. (2011). Learning analytics: Definitions, processes and potential. Retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

Fırat, M., & Yüzer, T. V. (2016). Learning analytics: Assessment of mass data in distance education. *International Journal on New Trends in Education and Their Implications, 7*(2), Article 01.

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text–based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*, 87–105.

Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education, 10*, 157–172.

Gasavic, D., Dawson, S., Mirriahi, N., & Long, P. D. (2015). Learning analytics–A growing field and community engagement. *Journal of Learning Analytics, 2*(1), 1–6.

Grau–Valldosera, J., & Minguillon, J. (2014). Rethinking dropout in online higher education: The case of the Universitat Oberta de Catalunya. *The International Review of Research in Open and Distance Learning, 15*(1), 290–308.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society, 15*(3), 42–57.

Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A measure of college student course engagement. *Journal of Educational Research, 98*(3), 184–192.

Hu, L., & Bentler, P. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling. Concepts, issues, and applications* (pp. 76–99). London: Sage Publication.

Hung, J–L., Hsu, Y–C., & Rice, K. (2012). Integrating data mining in program evaluation of K–12 online education. Educational Technology & Society, 15(3), 27–41.

Jaggars, S. S. (2014). Choosing between online and face–to–face courses. Community college student voices. *American Journal of Distance Education, 28*(1), 27–38.

Jo, I. H., Kim, D., & Yoon, M. (2015). Constructing proxy variables to measure adult learners' time management strategies in LMS. *Educational Technology & Society, 18*(3), 214–225.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC Horizon Report: 2016 Higher Education Edition*. Austin, Texas: The New Media Consortium.

Joksimovic, S., Gasevic, D., Kovanovic, V., Riecke, B. E., & Hatala, M. (2015). Social presence in online discussions as a process predictor of academic performance. *Journal of Computer Assisted Learning, 31*(6), 638–654.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guilford Press.

Knight, S., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, assessment, pedagogy: Where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23–47.

Koç, M. (2016). 2010–2015 yılları arasında öğrenme analitiklerinde araştırma eğilimleri [Research trends in learning analytics from 2010 to 2015]. *Proceedings of International Conference on Research in Education & Science (ICRES)*, 203–209.

Kong, S. C., Chan, T–W., Griffin, P., Hoppe, U., Huang, R., Kinshuk, ... Yu, S. (2014). E–learning in school education in the coming 10 years for developing 21st century skills: Critical research issues and policy implications. Educational Technology & Society, 17(1), 70–78.

Kupczynski, L., Ice, P., Wiesenmayer, R., & McCluskey, F. (2010). Student perceptions of the relationship between indicators of teaching presence and success in online courses. *Journal of Interactive Online Learning, 9*(1), 23–43.

Larusson, J. A., & White, B. (2014). *Learning analytics: From research to practice*. New York: Springer.

Leony, D., Munoz–Merino, P. J., Pardo, A., & Kloos, C. D. (2013). Provision of awareness of learners' emotions through visualizations in a computer interaction–based environment. *Expert Systems with Applications, 40*(13), 5093–5100.

Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 31–40. Retrieved from https://net.educause.edu/ir/library/pdf/ERM1151.pdf

Melero, J., Hernandez–Leo, D., Sun, J., Santos, P., & Blat, J. (2015). How was the activity? A visualization support for a case of location–based learning design. *British Journal of Educational Technology, 46*(2), 317–329.

Ng'ambi, D. (2013). Effective and ineffective uses of emerging technologies: Towards a transformative pedagogical model. *British Journal of Educational Technology, 44*(4), 652–661.

Pardo, A. (2014). Designing learning analytics experiences. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 15–38). New York: Springer.

Picher, P., & Ebner, M. (2015). Development of an information system to enhance students reading literacy. *International Journal of Emerging Technologies in Learning, 10*(3), 15–21.

Qvist, P., Kangasniemi, T., Palomaki, S., Seppanen, J., Joensuu, P., Natri, O., … Nordstrom, K. (2015). Design of virtual learning environments: Learning analytics and identification of affordances and barriers. International Journal of Engineering Pedagogy, 5(4), 64–75.

Rocca, K. A. (2010). Student participation in the college classroom: An extended multidisciplinary literature review. *Communication Education, 59*(2), 185–213.

Rovai, A. P. (2002). Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The Internet and Higher Education, 5*, 319–332.

Shea, P. (2006). A study of students' sense of learning community in online environments. *Journal of Asynchronous Learning Networks, 10*(1), 35–44.

Shea, P., Li, C. S., & Pickett, A. (2006). A study of teaching presence and student sense of learning community in fully online and web–enhanced college courses. *The Internet and Higher Education, 9*, 175–190.

Siemens, G., & Gasevic, D. (2012). Guest editorial–Learning and knowledge analytics. *Educational Technology & Society, 15*(3), 1–2.

Skylar, A. A. (2009). A comparison of asynchronous online text–based lectures and synchronous interactive web conferencing lectures. *Issues in Teacher Education, 18*(2), 69–84.

Spector, J. M. (2013). Emerging educational technologies and research directions. *Educational Technology & Society, 16*(2), 21–30.

Stiles, R., Jones, K. T., & Paradkar, V. (2011). Analytics rising: IT's role in informing higher education decisions. *ECAR Research Bulletin, 7*. Retrieved from https://net.educause.edu/ir/library/pdf/ERB1107.pdf

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn & Bacon.

Tabuenca, B., Kalz, M., Drachsler, H., & Specht, M. (2015). Time will tell: The role of mobile learning analytics in self–regulated learning. *Computers & Education, 89*, 53–74.

Taraghi, B., Saranti, A., Ebner, M., Müller, V., & Grobmann, A. (2015). Towards a learning–aware application guided by hierarchical classification of learner profiles. *Journal of Universal Computer Science, 21*(1), 93–109.

Tayebinik, M., & Puteh, M. (2013). Does greater participation in online courses lead to passing grade? An EFL learning context. *British Journal of Educational Technology, 44*(6), E199–E202.

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data–rich context. *Computers in Human Behavior, 47*, 157–167.

Teo, T. (2010). Using structural equation modeling (SEM) in educational technology research: Issues and guidelines. *British Journal of Educational Technology, 41*(6), E117–E119.

van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education, 79*, 28–39.

Williams, E. A., Duray, R., & Reddy, V. (2006). Teamwork orientation, group cohesiveness, and student learning: A study of the use of teams in online distance education. *Journal of Management Education, 30*, 592–616.

Yasmin, Dr. (2013). Application of the classification tree model in predicting learner dropout behavior in open and distance education. *Distance Education, 34*(2), 218–231.

Yıldız, E., & Bahçeci, F. (2014, September). Öğrenme *yönetim sistemlerinde kullanılan* öğrenme *analitikleri araçlarının incelenmesi* [Analyzing learning analytics tools used in learning management systems]. Paper session presented at the 8th International Computer and Instructional Technology Symposium, Edirne, Turkey.