Liu, L. (2011). Factors influencing students' preference to online learning: Development of an initial propensity model. *International Journal of Technology in Teaching and Learning*, 7(2), 93-108.

Factors Influencing Students' Preference to Online Learning: Development of an Initial Propensity Model

Leping Liu
University of Nevada, Reno

Online learning is influenced by a variety of factors, one of which is a student's preference to online learning. In the case when a course is only offered online, instructors may want to be aware of students' preference or any special needs for the purpose of creating an effective online learning environment. In current study, a set of factors are examined to determine the propensity of a student's preference to online learning. Logistic regressions are conducted and an initial propensity model is generated. The procedures to develop a propensity model are introduced. This model is then used for learner assessment in our online courses, and for further studies to reduce some nonrandom-grouping bias.

Keywords: online learning, propensity model, logistic regression, content analysis

INTRODUCTION

In the field of education, the term o*nline learning* is defined as learning that takes place partially or entirely over the Internet, including online delivery of course materials and instructions, interactive online learning activities, online communications, and online assessment (Cavanaugh, 2001; Maddux, Liu, Cummings, 2010). Online courses mostly are delivered through online course systems such as *Blackboard Learn* and a variety of free online learning platforms such as *Moodle* and some supporting cloud resources such as *Schoology*, *Edu 2.0*, or *Collaborize Classroom* (Cavus, Uzonboylu, & Ibrahim, 2007; Lai, Liu, Kiger, Jones, & Yuen, 2010; Li & Liu, 2011; Liu, 2010; Liu, Li, & Maddux, 2012). More and more online programs and online courses have been created in K-12 schools and universities in the United State. It is estimated and reported that there has been a 60 percent increase of the school districts in the States that have been offering online courses to students, (Zandberg & Lewis, 2008), and a 65 percent increase of students that have been participating in online learning (Means, Toyana, Murphy, Bakia, & Jones, 2010).

Leping Liu is Professor of Information Technology in Education and Statistics, in Counseling and Educational Psychology program, College of Education, University of Nevada, Reno. Please contact Dr. Leping Liu at liu@unr.edu

In online courses, a variety of factors could influence student learning, including (a) formats of course content delivery, (b) levels and types of online communications, (c) levels of computer anxiety, and (d) levels of students computing skills (Liu, 2006; Liu, D'Andrea, & Maddux, 2010; Sun, Lin, & Yu, 2008; Wallace & Clariana, 2000). The author's first-hand experiences have also revealed another influencing factor—a student's preference to online learning. In the case when a course is only offered online and all students have to take it online without any other choices, we as instructors will need to be aware of and deal with the differences between students who prefer online learning and who do not prefer, so that we could create an effective online learning environment to meet with the diversified needs of students. Furthermore, we also want to know if the preference to online learning could be predicted from some relevant factors.

The main purposes of current study are

1. to determine the set of factors that may influence the propensity of a student's preference to online learning,
2. to develop an initial propensity model that could be used to predict a student's preference of online learning, and
3. to demonstrate the procedures of developing a propensity model.

Further applications of this propensity model would be discussed, focusing on how it could be used to reduce sampling bias in studies of online learning where random grouping is difficult to obtain.

BACKGROUND INFORMATION

ONLINE LEARNING RESEARCH REVIEW

Online learning has become popular because it provides more flexible access to content and instruction at any time, from any place. Frequently, "the focus entails (a) increasing the availability of learning experiences for learners who cannot or choose not to attend traditional face-to-face offerings, (b) assembling and disseminating instructional content more cost-efficient, or (c) enabling instructors to handle more students while maintaining learning outcome quality that is equivalent to that of comparable face-to-face instruction" (Means, Toyana, Murphy, Bakia, & Jones, 2010, p. 1). Researchers and educators have conducted studies to examine (a) variables that influence the effectiveness of online learning (Aberson, Berger, Healy, & Romero, 2003; Beal, Kemper, Gardiner, & Woods, 2006; Golanics, & Nussbaum, 2008; Kock & Chatelain-Jardon, 2008), and (b) different perspectives of students about online learning and face-to-face learning (Aragon, Johnson, & Shaik, 2000; Bello, et.al., 2005; Campbell, Gibson, Hall, Richards, & Callery, 2008; Peterson, & Bond, 2004; Scheines, Leinhardt, Smith, & Cho, 2005; Sitzmann, Kraiger, Stewart, & Wisher, 2006).

In a meta-analysis conducted by the Center for Technology and Learning, U.S. Department of Education (Means et al., 2010), 176 studies have been analyzed on a set of critical variables of online learning. Findings suggest that (a) individualized instruction has positive impact on online learning (Grant & Courtoreille, 2007; Nguyen, 2007), (b) conditions in which learners have more control of their learning produce larger learning gains than do instructor-directed conditions (Cavus et al. 2007; Dinov, Sanchez, & Christou, 2008; Gao & Lehman, 2003; Zhang 2005), and (c) the effects of using different types of online simulation such as 2-dimension or 3-dimension images and animations are positive (Castaneda, 2008; Hibelink, 2007; Loar, 2007; Maki & Maki, 2002). In the designs of individualized instructions, self-controlled online learning conditions, and online simulations, instructors have also taken into consideration of students' needs, features, and preference to online learning formats.

In the literature review for this study, a content analysis was conducted on reference articles that examined the effectiveness of online learning, and a list of themes was generated (as shown in Figure 1), which have addressed the very common issues in the field of online teaching and learning.

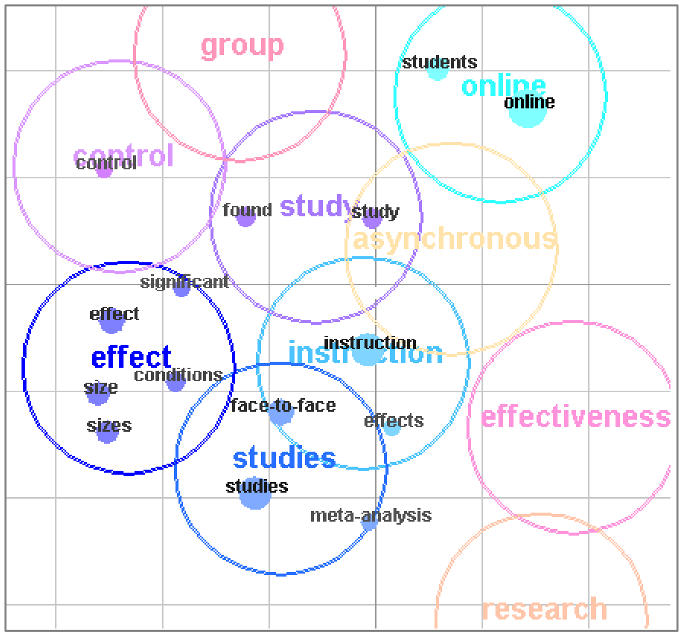

Figure 1. Main themes in online learning literature.

It is obvious that themes such as instructional design, online learning conditions, level of self-control, online group activities, and synchronous versus asynchronous communications are among those widely-studied themes. Most studies have also examined differences between online and face-to-face learning on a series of variables (Means et al., 2010).

STUDENTS' VIEW OF ONLINE LEARNING

The rationale of this study was also laid on students' perceptions, experiences, and views about their online learning. The author has taught an online course to teacher education students in online teaching and learning design. This course required students to participate a series of discussions focusing on the design of online learning, issuses and problems, and their individual lessons or experiences both as an online student and online instructor. The discussions generated over 400 messages. A content analysis was then conducted on the messages, and the main themes were obtained and mapped as shown in Figure 2.

Students' views on instructional design procedures, content design, online learning theories, and use of web technology were consistant with what was suggested in the literature. Furthermore, supprisingly, the rest themes were distributed more into areas of personal features or views, such as interest, enjoyment, anxiety, knowledge, believe, learning style, or time. Students' opinions on these themes clearly indicated the influence of individual differences on students' online learning.

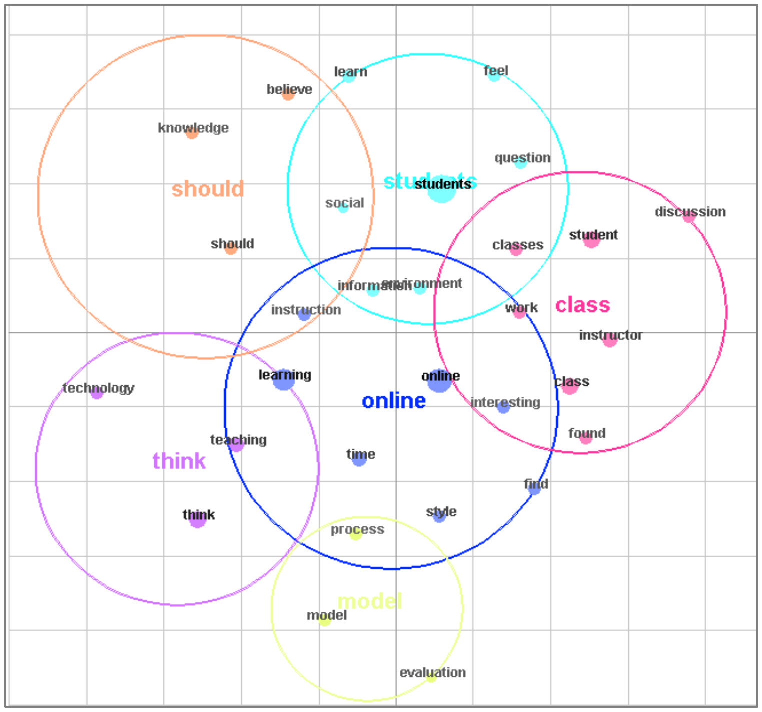

Figure 2. Factors influencing online learning.

A number of students expressed that they did prefer to face-to-face learning rather than online learning, and made an argument why it worked for them. This indeed brought an attention to us that in our online student porpulation, there might always be a group of traditional learners. Our online teaching may need to adjust to meet with the needs of those prefering online learning and those prefering face-to-face learning as well. Next, the literature review continued to demonstrate whether or how students' preference to course delivery format relates or influences their online learnng.

STUDENTS' PREFERENCE TO ONLINE LEARNING

Another content analysis on the literature reference articles that examine *students' preference to online learning* has generated a concept map (as shown in Figure 3) and a set of main themes. Literature suggests that students' preference to online learning has positively influenced their learning outcomes (Aragon, Johnson, & Shaik, 2002; Eom, Wen, & Ashill, 2006; Pawan, Paulus, Yalcin, & Chang, 2003; Wallace, 2003). Factors that relate to students preference to online learning or face-to-face learning include: students' psychological attitudes such as independence, creativity, tough-mindedness, sociability, risk-taking, stimulus-and sensation-seeking (Katz, 2002), students' readiness for online learning (Smith, 2005), students' learning styles and personality styles (Diaz, & Cartnal, 1999; Neuhauser, 2002; Sakagami & Kamba, 1997), students' preference in instructional methods (Butler & Pinto-Zipp, 2006), and students' preference in communication and interaction methods and styles (Northrup, 2002).

Obviously, findings from these studies provide a solid theoretical foundation for research in this area; and understanding the impact of those factors would definitely benefit the design of an effective online learning environment. However, realistically, when teaching an online course, an instructor always needs to deal with the differences or specific individual features in that particular class. The examinations on factors such as psychological attitudes, learning styles, and personality styles are time-consuming tasks;

it is not likely that the information could be finalized from any learner-assessment procedures conducted at the beginning of the semester. Therefore, it would be worthwhile to examine whether the more general and easily-obtained variables from student demographic data, such as major, gender, working status, years of education, online course experiences, technology preparation, communication styles, and pace of individual learning (Liu & D'Andrea, 2010; Liu & Maddux, 2008), could be factors that contribute to students' preference to online learning.

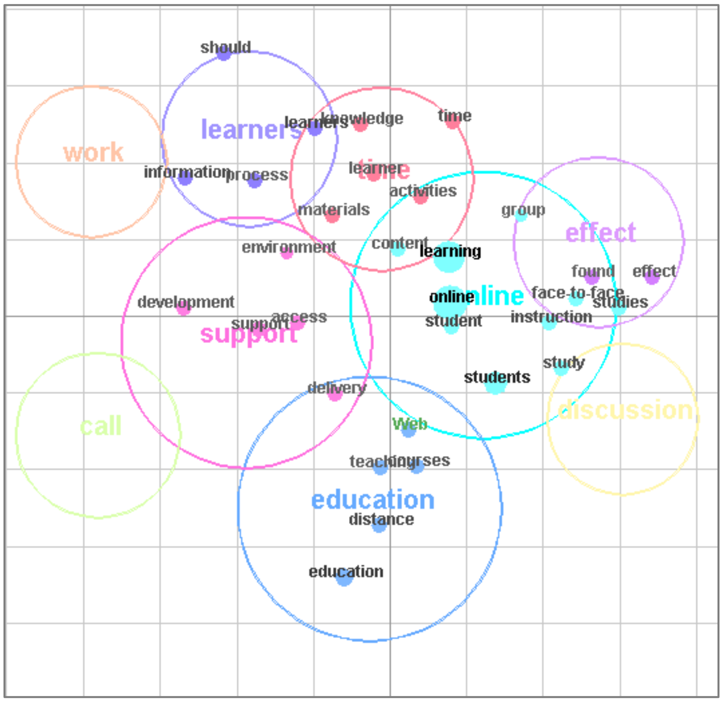

Figure 3. Factors related to students' preference to online learning.

To identify such influential variables in relation to students' preference to online learning, the methods of developing a logistic propensity model could be employed (Liu & D'Andrea, 2010; Liu & Maddux, 2005, 2008). Again, the main purpose of current study is to determine a set of critical variables and generate a model to predict the propensity that a student prefers to online learning. This propensity model could then be used as reference in the design of individualized instructions for online courses, and in online learning research to reduce the bias of non-random sampling.

RESEARCH QUESTIONS

Research questions examined in this study are:

1. Can the probability that a student prefers to online learning be predicted by any of these variables—major, gender, working status, online course experiences, technology preparation, communication styles, and pace of individual learning?
2. To what extent do the significant variables (if any from question 1) influence the probability of a student's preference to online learning?

METHODS

PARTICIPANTS

Participants in the study were 156 graduate students enrolled in six online courses offered from the college of education in a western university. Among them, 103 were females and 53 were males; 123 were education major and 33 were non-education major. They enrolled in the online courses because the courses were offered only online at the time they enrolled.

PROCEDURES AND MEASUREMENTS

A questionnaire was sent to students who were taking online graduate information technology courses at the beginning of semester. The questionnaire collected information on eight variables: (a) Preference of course delivery (PCD) – whether a student prefers to take an online course or a traditional face-to-face course; (b) Gender; (c) Technology skills – whether a student considers him/herself computer skillful or does not feel comfortable using computer technologies; (d) Online learning experiences –whether a student has taken at least one online course before or not taken any at all; (e) Working status – whether a student has a full time job or not; (f) Major – whether a student is majoring in Education or not; (g) Communication Style – whether a student feels individual communication with the instructor or class-wide communication as more helpful; and (h) Learning pace – whether a student feel him/herself learns better with self-paced learning or instructor controlled learning. The questionnaire was collected in the first week of the semester. Information was then coded as shown in Table 1, and entered to a database.

Table 1: Variable Coding

Variables	**Values**	
	1	**0**
(PCD) – Preference of Course Delivery (RV)	Online Learning	Face-to-Face Learning
(G) – Gender (EV)	Female	Male
(T) – Technology Skills (EV)	Skillful	Not Skillful
(O) – Previous Online Courses (EV)	Taken	Not Taken
(W) –Working Status (EV)	Full Time Students	Working
(M) – Major (EV)	Education	Non-Education
(C) – Communication Style (EV)	Individual	Course-wide
(L) – Learning-Pace (EV)	Self-Paced	Instructor-Controlled

Note: RV—Response Variable, EV—Explanatory Variable

DATA ANALYSIS

Logistic regression analyses were carried out to determine whether gender (G), technology skills (T), Online learning experiences (O), working status (W), major (M), communication style (C), and learning pace (L) could be used to predict a student's preference to course delivery (PCD). Logistic regression is a method of statistical modeling appropriate for categorical outcome variables. It describes the relationship between a categorical response variable and a set of explanatory variables. The response variable is usually dichotomous; typically the two outcomes are either "yes" or "no." The explanatory variables can be categorical or continuous.

The seven *explanatory variables* were: gender (*G*), technology skills (*T*), Online learning experiences (O), working status (*W*), major (M), communication style (C), learning pace (L) and the *response variable* was preference to course delivery (PCD). They were all coded with the values of 1 and 0, and the meanings of 1 or 0 are defined in Table 1. The assumptions of logistic regression were checked and no violations were found.

RESULTS

To determine the influential variables for the model, a logistic regression analysis was conducted to examine all seven explanatory variables. The results showed that two of the seven variables were not significant to the model: major (*Wald* $X^2 = 1.442$, $p = 0.23$) and communication style (*Wald* $X^2 = 2.988$, $p = 0.084$). Therefore these two variables were eliminated from the model in the next model examination. The five explanatory variables included in the next logistic regression analysis were: gender (*G*), technology skills (*T*), Online learning experiences (*O*), working status (*W*), learning pace (L).

Results from the second logistic regression showed that the second model with these five explanatory variables was significant ($X^2 = 66.428$, $p < 0.001$) and accounted for about 56% of the variation in the response variable ($R^2 = 0.563$), indicating that this model significantly predicts group membership. The Hosmer and Lemeshow Goodness-of-Fit Statistic of 14.807 ($p < 0.063$) was not significant, indicating that the hypothesis that the model provides a good fit of data should be accepted.

The receiver operating characteristic (ROC) curve, plotted from this model (see Figure 4.), rises quickly and the area under the curve is considerably large, indicating that this logistic regression model has relatively high predictive accuracy. This model can be used to predict a student's preference to online learning, as 53out of 74 scores (about 72%) of 0 (prefer to traditional face-to-face learning) and 62 out of 82 scores (about 76%) of 1 (prefer to online learning) were successfully predicted by the model.

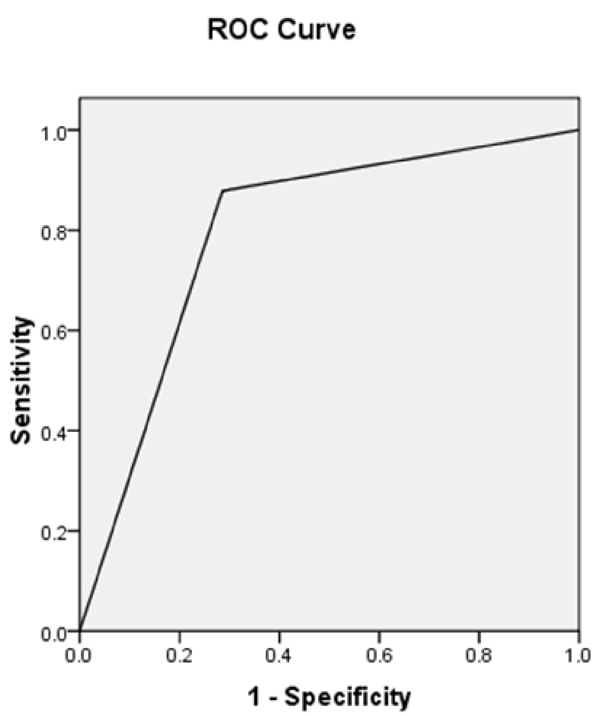

Figure 4. Receiver operating characteristic (ROC) curve.

Table 2: Logistic Regression Outputs

	DF	**Parameter Estimate**	**Standard Error**	**Wald Chi-Square**	***P***	**Odds Ratio**
(G)	1	-1.947	0.454	18.350	0.001	0.143
(T)	1	2.316	0.871	7.068	0.008	10.135
(O)	1	0.328	0.164	3.974	0.046	1.388
(W)	1	-1.089	0.415	6.892	0.009	0.336
(L)	1	1.522	0.643	5.603	0.018	4.583
Constant	1	-5.225	1.695	9.507	0.002	0.005

Response variable: preference to course delivery (PCD)
Explanatory variables: gender (G), technology skills (T), online learning experiences (O), working status (W), learning pace (L)

A significant Wald chi-square value for a given variable indicates that the variable is significantly related to the response variable. As shown in Table 2, the Wald chi-square values were significant for all five explanatory variables. Therefore, all five explanatory variables were included in the model equation. The Parameter Estimate generates the estimated coefficients of the fitted logistic regression model that are used to formulate the following logistic regression equation (1):

$$\text{logit}\ (\hat{p}) = -5.225 - 1.947(G) + 2.316(T) + 0.328(O) - 1.089(W) + 1.552(\mathrm{L}) \text{ ------ } (1)$$

The sign ($\hat{p}$) indicates an estimated probability value for the response variable to be 1, and logit represents *logit transformation* of the event probability.

An estimated coefficient indicates the contribution that particular explanatory variable makes to the possibility of the response variable being 1. For example, when the variable *T* (Technology skills) is 1 (that is, when the student considers him/herself technology-skillful), the logit transformation of event probability (that he/she would prefer to take an online course) increases by 2.316.

If the odds ratio for a given variable is larger than 1, the probability of the response variable being 1 increases. For example, the odds ratio for variable *T (technology skills*) is 10.135. This means that a student would be 10.135 times more likely to take an online course if he or she is self-rated as technology skillful, compared to those who rate themselves as not comfortable with technologies.

If the odds ratio is smaller than 1, the probability of the response variable being 1 decreases. For example, the odds ratio for variable *W* (working status) is 0.336. This means that a full-time student would be 0.336 times (less) likely to take an online course, compared to those working-students. In another word, it could also be understood as that a working-student would be more likely to take an online course than a full-time student.

In summary, the results indicate that a student is more likely to prefer online learning or take an online course if he/she feels comfortable with technologies, has taken some online courses before, works full time, and learns better with more self-control on the pace of learning. Also, in this study, gender does significantly influence a student's preferences. Furthermore, the logistic regression equation presented above has clearly provided answers to the research questions, and clarified the five significant explanatory variables and described the influence of each variable to the probability of a student's preference to online learning.

THE INITIAL PROPENSITY MODEL

One purpose of this study was to develop an initial propensity model to predict a student's preference of online learning, and to demonstrate the procedures of developing a propensity model. Before the discussion of findings from the data analysis results, I will first summarize the model and briefly introduce the basics of *propensity model* and *propensity scores*, using the initial model developed from this study as an example.

MODEL FUNCTION

Results and relationships produced from the logistic regression data analysis can be summarized into the following model function (2) in Figure 5.

$$\mathbf{P\ (PCD) = \mathit{f}\ [G, T, O, W, L]} \quad \text{--------------------------------------} \quad (2)$$

Where:

PCD = *Preference to Course Delivery*	**P (PCD) =** *Probability of PCD*
G = *Gender*	**T** = *Technology Skills*
O = *Online Learning Experiences*	**L** = *Learning Pace Control*
W = *Working Status*	***f*** **[…]** indicates "a function of …"

Figure 5. Model function.

Model function (2) reads "the probability of students' preference to course delivery is a function of all five variables – gender, technology skills, online learning experience, working status, and learning pace." It exhibits the relations between the group of explanatory variables and the response variable. Logistic regression equation (1) in the "Results" section is the concrete model that describes all specific predictive relations or influences.

PROPENSITY MODEL AND PROPENSITY SCORES

Logistic regression equation (1), the *PCD propensity model*, is an initial propensity model to predict the probability that a student prefers to online learning. It can be used to calculate propensity scores with the data from any other samples of online students.

In the literature, propensity scores are commonly used in research design to reduce selection bias by equating groups based on these covariates (Guo & Fraser, 2010), in the case when random sampling is not possible. Researchers have defined the term *propensity score* as the probability of a unit (e.g., person, classroom, school) being assigned to a particular condition in a study given a set of known covariates (Rosenbaum & Rubin, 1983, 1984, 1985).

For example, in the *PCD propensity model*, the set of known covariates are the five explanatory variables. A propensity score calculated from the model equation indicates the probability of a student's PCD (preference to course delivery) to be online. Using the propensity scores calculated from a particular group of students, an instructor can identify students' preference to online learning, which would be of reference when deigning individualized online instructions. Using propensity scores, a researcher can create comparison groups with matching methods. Relative to matching directly on the covariates, propensity score matching has the advantage of reducing the dimensionality of matching to a single dimension (Blackford, 2009; Segal, et al., 2007).

PROCEDURES OF MODEL DEVELOPMENT

Development of a propensity model is based on a series of mathematical theories and procedures (Guo & Fraser, 2010). Emphasis of this article is laid on the aspect of its application. The following are the main procedures to develop a propensity model:

1. *Identifying the treatment condition*.
   This is the response variable in the logistic regression analysis. For example, such treatment condition can be online versus face-to-face, likely or unlikely to major Engineering, or likely versus unlikely to have cancer.
2. *Determining a set of known covariates.*
   These should be all possible variables relevant to the possibility to be assigned to one of the two treatment conditions. Generally, the selection of this set of variables is based on research literature and the researcher's practical experiences. In this study, for example, the seven explanatory variables tested first. They will be the explanatory variables in the logistic regression analysis.
3. *Determining variables to be included in the model.*
   These critical variables are determined through a series of logistic regression analyses. This is also a process of model selection:
   a. Use all covariates determined in procedure 2 as the explanatory variables in the first logistic regression analysis.
   b. Any variables if not significant from the first run should be excluded.
   c. Run another logistic regression only with variables that are significant in the first run.
   d. Repeat procedures b and c, and try different combinations of the variables, until reaching a model consist of variables that are all significant.
4. *Finalizing the model.*
   Use the logistic regression results from the final model from above procedure 3-d (similar to the results in Table 2), and generate the model expressed as in the logistic regression equation (1). This is the propensity model equation that a researcher can use to calculate propensity scores.

The procedures of model development described above can be used specifically in developing a propensity model, or generally in developing a multiple predictive model (Liu, Maddux, & Johnson, 2004).

DISCUSSIONS AND CONCLUSIONS

Addressing back to the purposes of the study, the influential variables are determined that can be used to predict the propensity of a student's preference to online learning, an initial propensity model is formulated, and procedures of model development and basic applications of the propensity model are presented.

RESEARCH QUESTION ONE

Can the probability that a student prefers to online learning be predicted by any of the variables—major, gender, working status, online course experiences, technology preparation, communication styles, and pace of individual learning?

This study has found that the probability that a student prefers to online learning can be predicted by five variables – gender, working status, online course experiences, technology preparation, and pace of individual learning.

The predicted probability of a student's PCD (preference to course delivery) by a combination of the five variables will provide more insights for an instructor to understand the detailed preferences or special needs of the student. When the probabilities of PCD from an entire class are predicted, the distribution of the probabilities will provide a thorough understanding of students as a group about their preferences. For example, according to this distribution of the probabilities (that is, knowing the extent to which each student prefers online learning), an instructor is able to adjust the design of individualized instructions (Grant & Courtoreille, 2007), create online learning environment in which learners have more control of their learning (Dinov, Sanchez, & Christou, 2008), or provide different types of online simulation such as 2-dimension or 3-dimension images and animations (Castaneda, 2008). With information of detailed distribution of the probabilities, instructional decisions could be made as close as possible to meet the specific needs of the whole group.

RESEARCH QUESTION TWO

To what extent do the significant variables (if any from question 1) influence the probability of a student's preference to online learning?

The coefficients in the logistic regression equation explained the extent each significant variables influences the probability of a student's preference to online learning.

From the perspective of learner assessment, this study took a multiple dimensional approach and developed a propensity model to produce a score (the propensity score) that combines the influences from all five explanatory variables. A propensity score calculated with combined data from multiple dimensions would provide more accurate information than that from a single dimension. In this study, the *multiple-dimension* is the five significant variables. Data of the five variables are used to calculate the probability that a student prefers online learning. A single dimension data could be that from a "yes" or "no" question; however, it does not provide the extent to which a student prefers online learning. This explains the advantage of using a propensity score to obtain multiple dimension information over the use of that from a single dimension question.

USING PROPENSITY SCORES FOR RESEASRCH SAMPLING

Originally, propensity scores are used in research when random sampling is not likely to be conducted, for example, in medical or biometric research (Ye & Kaskutas, 2008), to estimate causal effects from large sets of existing data (Jenkins et al., 2007) and nonexperimental longitudinal data (Haviland et al., 2008), or to reduce the samplings bias of Internet surveys (Lee & Valliant, 2009).

In most studies, propensity scores are used to create comparison groups with *matching methods*. For example, in a study in which data were collected from a convenient sample of two classes, one class served as experimental group, and the other as the control group. A propensity model can be used to calculate propensity scores for both classes, obtain two sets of propensity scores—the probabilities that a student is likely to be assigned to the treatment. The two sets of propensity scores will be the two sets of probability distribution. With individual matching or tier-matching (Guo & Fraser, 2010), the bias of this non-random grouping could be reduced.

Another application of propensity scores is to use the method of *propensity score dividing*. For example, in an online class, the researcher intended to examine a new method with students who prefer to use new technology, comparing with those who do not prefer. A propensity model can be used to calculate the propensity scores from each one in the class, obtaining a distribution of the probabilities that students prefer to use new technology. From this distribution, the researcher could choose the scores from the top 5% and bottom 5%, formulating two groups. Students with the top 5% probabilities are the group that most likely to prefer the new technology, those with the bottom 5% probability would be the group that least likely to use new technology. Of course this method would involve *dividing* method and more *multiple matching* methods (Rosenbaum & Rubin, 1983, 1984, 1985), but it does provide a sampling method to reduce the bias caused by nonrandom samples.

CONCLUSIONS AND FURTHER STUDIES

The data analyses results, formulation of the PCD propensity model, and the application of a propensity model in general could lead the study to three conclusions.

First, any factor related to student learning (e.g., a student preference to online learning, in this study) would be influenced by influential variables from multiple dimensions. In the design of online instructions, materials, activities, and even the entire online course, the influences from all the variables need to be considered, including (a) how the influences from each variable could interact with one another, (b) whether of how the influences from each variable should be weighted or that from all the variables be balanced, and (c) whether or what other possible variables could have additional influences on student learning. The research agenda within this scope includes an endless list of study topics.

Second, variables that influence online learning can be sorted into levels, including variables that have direct and indirect influences to online learning. In the context of current study, effectiveness of online learning is influenced by different perspectives students have about online and face-to-face learning (Aragon, Johnson, & Shaik, 2002; Bello, et.al., 2005; Campbell, et al., 2008). Student preference to online learning should be at this first level, directly influencing online learning. The five influential variable determined in this study are at the second level, directly influencing student preference to course delivery format, but indirectly influencing online learning. These leveled, direct/indirect relationships between all influential variables and online learning are also critical to the design of an effective online learning environment.

Third, model selection is a key procedure not only in the development of a propensity model, but also in the development of an effective online learning environment. For example, in the design of an online course, the "selection" of online materials, activities, communication styles, or assessment strategies would all involve the determination of (a) what is the appropriate choice to the particular group of learners, (b) what is the best combination of all the possible choices, and (c) what are best choices that have positive impact, directly or indirectly, on online learning. Model selection is more a mathematical procedure, and the *best-choice* design selection is more a combined procedures.

The current study started with the examination of seven related variables and ended with five significant ones included in the initial propensity model. In further studies, to continue building a more solid model, examinations may focus on more possible variables from students in different major areas such as science, engineering, arts and others. Moreover, to develop a propensity model of broader application scope, data from some national educational database could be used. It is the author's hope that this initial step could be of reference to the work of other educators and researchers.

REFERENCES

Aberson, C. L., Berger, D. E., Healy, M. R., & Romero, V. L. (2003). Evaluation of an interactive tutorial for teaching hypothesis testing concepts. *Teaching of Psychology, 30*(1), 75–78.

Aragon, S., Johnson, S., & Shaik, N. (2000). The influence of learning style preferences on student success in online vs. face-to-face environments. In: *Proceedings of WebNet World Conference on the WWW and Internet 2000* (pp. 17-22). Chesapeake, VA: AACE.

Aragon, S. R., Johnson, S. D., & Shaik, N. (2002). The influence of learning style preferences on student success in online versus face-to-face environments. *The American Journal of Distance Education, 16*(4), 227-244.

Beal, T., Kemper, K. J., Gardiner, P., & Woods, C. (2006). Long-term impact of four different strategies for delivering an online curriculum about herbs and other dietary supplements. *BMC Medical Education,* 6, 39.

Bello, G., Pennisi, M. A., Maviglia, R., Maggiore, S. M., Bocci, M. G., Montini, L., & Antonelli, M. (2005). Online vs. live methods for teaching difficult airway management to anesthesiology residents. *Intensive Care Medicine, 31*(4), 547–552.

Blackford, J. U. (2009). Propensity scores: Method for matching on multiple variables in down syndrome research. *Intellectual and Development disabilities, 47*(5), 348-357.

Butler, T. J., & Pinto-Zipp, G. (2006). Students' learning styles and their preferences for online instructional methods, *Journal of Educational Technology Systems,* 34(2), 199-221.

Campbell, M., Gibson, W., Hall, A., Richards, D., & Callery, P. (2008). Online vs. face-to-face discussion in a web-based research methods course for postgraduate nursing students: A quasi-experimental study. *International Journal of Nursing Studies, 45*(5), 750–759.

Castaneda, R. (2008). The impact of computer-based simulation within an instructional sequence on learner performance in a web-based environment. PhD diss., Arizona State University, Tempe.

Cavanaugh, C. (2001). The effectiveness of interactive distance education technologies in K–12 learning: A meta-analysis. *International Journal of Educational Telecommunications, 7*(1), 73–78.

Cavus, N., Uzonboylu, H., &Ibrahim, D. (2007). Assessing the success rate of students using a learning management system together with a collaborative tool in web-based teaching of programming languages. *Journal of Educational Computing Research, 36* (3), 301–321.

Diaz, D. P., & Cartnal, R. B. (1999). Students' learning styles in two classes: Online distance learning and equivalent on-campus. *College Teaching*, *47*(4), 130-135.

Dinov, I. D., Sanchez, J., & Christou, N. (2008). Pedagogical utilization and assessment of the statistic online computational resource in introductory probability and statistics courses. *Computers & Education, 50*(1), 284–300.

Eom, S. B., Wen, J., & Ashill, N. (2006). The determinants of students' perceived learning outcomes and satisfaction in university online education: An empirical investigation. *Journal of Innovative Education, 4*(2), 215-235.

Gao, T., & Lehman, J. D. (2003). The effects of different levels of interaction on the achievement and motivational perceptions of college students in a web-based learning environment. *Journal of Interactive Learning Research, 14*(4), 367–386.

Golanics, J. D., & Nussbaum, E. M. (2008). Enhancing online collaborative argumentation through question elaboration and goal instructions. *Journal of Computer Assisted Learning, 24*(3), 167–180.

Grant, L. K., & Courtoreille, M. (2007). Comparison of fixed-item and response-sensitive versions of an online tutorial. *Psychological Record*, *57* (2), 265–272.

Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Los Angeles, CA: Sage Publications.

Haviland, A., Nagin, D. S., Rosenbaum, P. R., & Tremblay, R. E. (2008). Combining group-based trajectory modeling and propensity score matching for causal inferences in nonexperimental longitudinal data. *Developmental Psychology, 44*(2), 422-436.

Hilbelink, A. J. (2007). The effectiveness and user perception of 3-dimensional digital human anatomy in an online undergraduate anatomy laboratory. PhD diss. dissertation, University of South Florida, Orlando.

Jenkins, P., Earle-Richardson, G., Burdick, P., & May, J. (2007). Handling nonresponse in surveys: Analytic corrections compared with converting nonresponders. *American Journal of Epidemiology, 167*(3), 369-374.

Katz, Y. J. (2002). Attitudes affecting college students' preferences for distance learning. *Journal of Computer Assisted Learning, 18*(1), 2-9.

Kock, N., & Chatelain-Jardon, R. C. J. (2008). An experimental study of simulated web-based threats and their impact on knowledge communication effectiveness. *IEEE Transactions on Professional Communication,* 51, 183–197.

Lai, F. Q., Liu, L., Kiger, S., Jones, & M., Yuen, S. (2010). Key factors to a successful online program. In C. Crawford et al. (Eds.), Proceedings of Society for Information Technology & Teacher Education International Conference 2010 (pp. 606-608). Chesapeake, VA: AACE.

Lee, S., & Valliant, R, (2009). Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment. *Sociological Methods and Research, 37*(3), 319-343.

Li. W., & Liu, L., (2011). The effects of the instructional video on pre-service teachers' technology learning in an online environment. In V. Wang, (ed.), *Encyclopedia of E-Leadership, counseling and Training* (pp. 656-665). Hershey, PA: IGI Global.

Liu, L. (2006). Communication design and team building in teacher education web based courses. In C. Crawford, R. Carlsen, I. Gibson, K. McFerrin, J. Price, R. Weber & D. A. Willis (Eds.), *Technology & Teacher Education Annual 2006* (pp. 2594-2599). Charlottesville, VA: AACE.

Liu, L. (2010). Assessment of online master programs: Procedures, management, and outcomes. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010* (pp. 176-182). Chesapeake, VA: AACE.

Liu, L., D'Andrea, L. & Maddux, C. (2010). An online course in assessing technology integration: Design, implementation, and issues. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010* (pp. 170-175). Chesapeake, VA:AACE.

Liu, L., & D'Andrea, L. (2010). Initial stages to create online graduate communities: Assessment and development. In V. Wang, (ed.), *Encyclopedia of information communication technologies and adult education integration* (pp. 911-926). Hershey, PA: IGI Global.

Liu, L., Li, W., & Maddux, C. (2012). Prepare teacher education students to use cloud resources: Evaluation, design, and integration. Paper accepted by the *Society for Information Technology & Teacher Education International Conference 2012.* March, 5-9, Austin, TX.

Liu, L., & Maddux, C. (2008). Web 2.0 articles: Content analysis and a statistical model to predict recognition of the need for new instructional design strategies. *Computers in the Schools*, *25*(3/4), 314-328.

Liu, L., & Maddux, C. (2005). Influences of course design on student evaluations: An initial logistic prediction model. *Journal of Excellence in College Teaching, 16*(1),125-148.

Liu, L., Maddux, C., & Johnson, L. (2004). Computer attitude and achievement: Is time an intermediate variable? *Journal of Technology and Teacher Education,12*(4), 593-607.

Loar, R. S. (2007). The impact of a computer simulated case study on nurse practitioner students' declarative knowledge and clinical performance. PhD diss., University of Illinois at Urbana-Champaign.

Maddux, C., Liu, L., & Cummings, R. (2010). A totally online university class in statistics for teachers: Aids and cautions. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010* (pp. 699-704). Chesapeake, VA: AACE.

Maki, W. S., & Maki, R. H. (2002). Multimedia comprehension skill predicts differential outcomes of web-based and lecture courses. *Journal of Experimental Psychology: Applied, 8*(2), 85–98.

Means, P., Toyana, Y., Murphy, R., Bakia, M., & Jones, K. (2010). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. Washington, D.C.: Office of Planning, Evaluation, and Policy Development, Center for Technology in Learning, U.S. Department of Education.

Neuhauser, C. (2002). Learning style and effectiveness of online and face-to-face instruction. *American Journal of Distance Education,16*(2), 99-113.

Nguyen, F. (2007). The effect of an electronic performance support system and training as performance interventions. PhD diss., Arizona State University, Tempe.

Northrup, P. T. (2002). Online learners' preferences for interaction. *Quarterly Review of Distance Education, 3*(2), 219-226.

Pawan, F., Paulus, T. M., Yalcin, S., & Chang, C. F. (2003). Online learning: patterns of engagement and interaction among in-service teachers. *Language Learning and Technology, 7*(3), 119-140.

Peterson, C. L., & Bond, N. (2004). Online compared to face-to-face teacher preparation for learning standards-based planning skills. *Journal of Research on Technology in Education, 36*(4), 345–361.

Rosenbaum, P. R., Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Rosenbaum, P. R., Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistics Association, 79*(387), 516-524.

Rosenbaum, P. R., Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistics, 39*(1), 33-38.

Sakagami, H., & Kamba, T. (1997). Learning personal preferences on online articles from user behaviors. *Computer Networks and ISDN Systems, 29*(8), 1447-1455.

Scheines, R., Leinhardt, G., Smith, J., & Cho, K. (2005). Replacing lecture with web-based course materials. *Journal of Educational Computing Research, 32*(1), 1–25.

Segal, J. B., Griswold, M., Achy-Brou, A., Herbert, R., Bass, E. B., Dy, S. M., Millman, A. E., Wu, A. W., & Frangakis, C. E. (2007). Using propensity scores subclassification to estimate effects of longitudinal treatments: An example using a new diabetes medication. *Medical Care, 45*(10), S149-S157.

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology*, 59, 623–664.

Smith, P. (2005). Learning preference and readiness for online learning. *Educational Psychology: An International Journal of Experimental Educational Psychology, 25*(1), 3-12.

Sun, K., Lin, Y., & Yu, C. (2008). A study on learning effect among different learning styles in a web-based lab of science for elementary school students. *Computers & Education, 50*(4), 1411–1422.

Wallace, P. E., & Clariana, R. B. (2000). Achievement predictors for a computer-applications module delivered online. *Journal of Information Systems Education, 11*(1/2), 13–18.

Wallace, R. M. (2003). Online learning in higher education: A review of research on interactions among teachers and students. *Education, Communication & Information, 3*(2), 241-280.

Ye, Y., & Kaskutas, L. A. (2008). Using propensity scores to adjust for selection bias when assessing effectiveness of alcoholics anonymous in observational studies. *Drug and Alcohol Dependence, 104*(1), 56-64.

Zandberg, I., & Lewis, L. (2008). Technology-based distance education courses for public elementary and secondary school students: 2002-03 and 2004-05. (NCES 2008-08). Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Zhang, D. (2005). Interactive multimedia-based e-learning: A study of effectiveness. *American Journal of Distance Education, 19*(3), 149–162.