3-25-2017

Applying the Technology Acceptance Model (TAM) to Automatic Grading Technology for Large Projects

Christopher Kreider
Georgia Institute of Technology, chriskreider@gatech.edu

APPLYING THE TECHNOLOGY ACCEPTANCE MODEL (TAM) TO AUTOMATIC GRADING TECHNOLOGY FOR LARGE PROJECTS

Christopher Kreider
Georgia Institute of Technology
chriskreider@gatech.edu

ABSTRACT

Autograding technology, a form of Computer Based Assessment (CBA), should allow course enrollments to grow without reducing the number of exercises, however, these gains are not expected to be immune to problems with adoption. This study utilizes the Technology Acceptance Model (TAM) to explore student and staff perceptions of autograding technology. This study extends prior research by specifically exploring the phenomena in the context of large projects. Our study explores the perceptions of 128 students and course staff in an online master degree program in computer science at a large public university. Our research design was chosen to leverage existing theories while also providing findings that will enable practitioners to apply them to their decision making regarding autograding technology. We find that perceived usefulness was significantly correlated with behavioral intention for both students and staff, leading to our hypotheses being supported and partially supported. Additionally, we find that perceived ease of use is only significantly correlated with student's intentions, and does not apply to course staff.

Keywords

Educational Technology, Technology Acceptance Model, Autograding, Adoption

INTRODUCTION

Autograding technology is expected to reduce the amount of effort on course staff as course enrollments continue to grow without reducing the number of exercises (Ihantola et al. 2010), but some students find it is excessively strict and unforgiving (Wilcox 2015). The Technology Acceptance Model (TAM) has shown that Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) are better predictors of adoption than many measures of usability which are more subjective (Davis 1989). Specifically, the effectiveness of autograding technology, a type of Computer Based Assessment (CBA), is expected to depend on the students acceptance (Terzis and Economides 2011). Additionally, this technology seems to have a high potential to increase performance in a variety of measures, specifically institutional costs and student feedback, but is not immune to barriers to achieving these gains due to poor adoption. In fact, it is common for businesses to experience problems with the introduction of new systems into their business processes. A significant portion of projects run into delays, budget shortcomings, and failure to meet functional requirements (Schepers and Wetzels 2007). Additionally, and despite many advances in technological capabilities, many organizations still experience underutilized systems (Venkatesh and Davis 2000).

This lack of utilization is unsurprising in the context that information technology, in general, has been heralded as a way to increase performance, but these gains are reliant on users willingness to accept and use the available systems, commonly known as adoption (Davis 1989). Adoption is considered a "broad area of inquiry"(Venkatesh et al. 2003:427), which makes it applicable to a wide variety of technology usages. It should therefore be important to explore adoption in educational environments where a student's or instructor's failure to accept a system can offset gains in performance the technology was intended to provide, such as increased student experiences/outcomes, lower costs, and increased quality.

Specifically, TAM has been explored in a variety of contexts, including educational technology (Atif et al. 2015; Cheung and Vogel 2013; Edmunds et al. 2012; Persico et al. 2014). In the context of adoption and computer based assessment, it is unclear as to the scale of the assessment with respect to the student's educational progress. For example, an automatic quiz grader that accounts for a small portion of a student's final grade (< 1%) has different implications when compared to an automatic project grader that accounts for a large portion of a student's final grade (≥15%). Using a simple survey based on the basic TAM model, we explore perceptions of autograding technology in a large online master's degree program in computer science at a large public university. Based on this, we seek to answer the following research questions.

RQ1: *How do perceived ease of use and perceived usefulness of autograding technology used for large project grading influence student's desire to use an autograder over a traditional grader?*

RQ2: *How do perceived ease of use and perceived usefulness of autograding technology used for large project grading influence an instructor's (professor/TA) desire to use an autograder over a traditional grader?*

The rest of this paper will be arranged as follows. We will first explore the extant literature relevant to the educational technology, the technology acceptance model (TAM), and Computer Based Assessment (CBA). We will then discuss our research model and hypotheses, our methodology, followed by our results, limitations and conclusions.

LITERATURE REVIEW

Educational Technology

In the context of student experience/performance, a meta-analysis of 45 studies exploring online and blended learning environments found that purely online environments were equivalent to face-to-face instruction in terms of effectiveness (Means et al. 2013). Additionally, the results of a 2013 survey recapping 10 years of online education in the United States found that only 77 percent of academic leaders rated online learning outcomes as the same or superior when compared to face-to-face alternatives. While these numbers have increased from prior years, a gap still exists (Allen and Seaman 2013). In a meta-analysis by Bernard et al. (2004) comparing distance to classroom based education of over 2,200 abstracts, the researchers found somewhat conflicting results. On one hand, they found that in many cases, distance education applications were viewed more positively, had higher retention rates, and better achievement results when compared to their classroom based counterparts. On the other hand, however, a substantial number of distance education programs were also much worse on the same three measures. These findings lead them to conclude that it would be incorrect to state that distance education was better, worse or even equal on the three measures explored. The major caveat of this study, however, was that the literature was found to be "severely wanted in terms of depth of reporting (Bernard et al. 2004:407)."

In the context of other factors such as financial gains, little to no research has shown that the implementation of online educational technologies has reduced the level of faculty involvement, and thus the cost of instruction (Green and Gilbert 1995), with to the contrary, an expectation that online courses require more time and effort to teach than face-to-face courses (Allen and Seaman 2013). In the context of education quality, some critics doubt that it possesses the capability to replace synchronous, face to face learning (Latchman et al. 1999), with some even considering it a "risk" that Massively Open Online Courses (MOOCs) will be considered a sufficient substitute for more traditional methods in education (Dasarathy et al. 2014).

Given the gains in performance that information technology is expected to enable, and the specific criticisms of it in educational technology that are readily available in the academic literature, one may question where this discrepancy occurs? Additionally, this question becomes more interesting in light of the conflicting results when comparing online to traditional education, with some studies finding that they are comparable (Means et al. 2013), while others effectively saying it is just different (Bernard et al. 2004). In the conclusion of a 2016 meta-analysis of technology-supported interaction in post-secondary education, the authors remind us that, "Technology is just a tool and it behooves us to learn to employ it effectively" (Borokhovski et al. 2016, p. 24:24). In this case, we argue that exploring barriers to adoption is just one aspect of learning to employ educational technology efficiently, and may help to make any potential benefits of online education more clear. One of the better and more well-known models of technology adoption is the Technology Acceptance Model (TAM), as proposed by Davis (Davis 1989; Park 2009; Schepers and Wetzels 2007) .

Technology Acceptance Model

By applying TAM, decision makers are better able to explore barriers to adoption, and make predictions about adoption when compared to other measures of usability. The primary factors of Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) have been shown to be better predictors of adoption than many subjective measures (Davis 1989). TAM has been applied educational technology in a variety of contexts (Atif et al. 2015; Cheung and Vogel 2013; Edmunds et al. 2012; Persico et al. 2014) and specifically to Computer Based Assessment (CBA) (Terzis and Economides 2011). In the study by Terzis and Economides (2011), the items in their survey fail to explore the difference between large and small projects that are the result of what the autograder is grading. From the common sense perspective, we consider this distinction interesting whereas problems with an autograder grading a trivial quiz ≤1% of the final grade are much different when grading non-trivial work in terms of time/effort/course performance. These differences and the relative cost of poor adoption is also likely to increase as course sizes increase, especially with the growth of Massively Open Online Course (MOOC) style classes.

From a course staff perspective, being able to effectively integrate autograding technology into a courses structure, including being able to interpret results and respond to students requesting clarification is expected to be important. In a study specifically applying TAM to faculty, Gibson et al. (2008) explore adoption of online education, finding that perceived usefulness was the most powerful predictor with perceived ease of use providing little advances in the predictive power of the model.

RESEARCH MODEL AND HYPOTHESES

Our research model is based on the basic Technology Acceptance Model (TAM) where Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) are predictors of Behavioral Intention (BI) (Davis 1989). Additionally, we chose to explore both student and course staff perceptions as literature indicated that there may be a difference, specifically with respect to PEOU (Gibson et al. 2008). We believed this to be feasible due to the large class sizes and the large number of course staff available for sampling.

Hypothesis

Our first two hypothesis explore student perceptions. Autograding technology is expected to reduce the amount of effort as course enrollments continue to grow without requiring staff to reduce the number of exercises (Ihantola et al. 2010), but in some cases, students have found it lacking in usability being excessively strict and unforgiving (Wilcox 2015), and its effectiveness is expected to depend on student acceptance (Terzis and Economides 2011). Specifically, large projects may introduce additional complexities pertaining to student's grade, and thus performance, in a course. The consequences of a problems with an autograder for a trivial quiz that may account for ≤ 1% of a final grade are much different than those associated with large projects that represent a large portion of a student's effort, and similarly, grade. Thus, we expect that PEOU and PU will be positively correlated BI.

H1: PEOU will be positively correlated with behavioral intention for students using the autograder for large projects

H2: PU will be positively correlated with behavioral intention for students using the autograder for large projects

The second two hypotheses specifically explore course staff, where one would expect that staff would be more likely to prefer autograding technology over human graders if they expect it to be useful and easy to use. However, in a study by Gibson et al. (2008) specifically exploring faculty adoption of online education, they found that perceived usefulness was the most powerful predictor with perceived ease of use providing little advances in the predictive power of the model. While the context of this study is different than ours, it provides a prediction when applying the technology acceptance model to educational technology with respect to course staff. We find the conclusion that PEOU added little to the models explanatory model interesting, and our exploration of this has the potential to replicate these past findings.

H3: PEOU will not be significantly correlated with behavioral intention for course staff using the autograder for large projects

H4: PU will be positively correlated with behavioral intention for course staff using the autograder for large projects

METHODOLOGY

The data was gathered to test the hypothesis through the distribution of a survey. The survey was issued to two online graduate level classes, as well as a private Google+group for current and past students in the Online Masters of Science in Computer Science (OMSCS) program at Georgia Tech. The survey was issued with an opening description of a sample project, describing the project size and scope and how the autograder would be used, and is shown in figure 1 below.

Student Perceptions of Autograder Adoption

This survey is exploring adoption of autograding technology in online educational environments. Specifically, this survey is referring to the use of an "autograder" to provide grades for a non-trivial final programming project, such as one requiring many weeks of work and accounting for 15%-20% of your final grade.

Your projects code would be submitted to the autograder, which would then execute your code on an external server. The final grade would be primarily based on the results of the autograder.

Figure 1. Survey Introduction

Some of the students may have received a very small amount of course credit, but due to the anonymous nature of the survey, we were unable to determine which respondents these may have been. The survey had a total of 128 respondents, of which, 105 identified as students and 23 identified as part of course staff. The 10 survey items were developed informally based on a brief exploration of an online message board for a course that relied heavily on autograders for large projects in the class, and are listed below in Table1. The survey items were developed using the TAM constructs as a base, but were specifically worded to be relevant to the technology being studied. The items were presented as a 7 point Likert scale with the ends labeled appropriately based on each question.

Question	**Factor**
I expect that using an autograder will allow me to get helpful feedback on a final project	PU1
I expect that using an autograder will allow me to get the most possible points on a final project	PU2
I expect that using an autograder will provide more consistent feedback than a human grader on a final project	PU3
I expect that an autograder will provide easy to understand feedback to help me understand what mistakes I made	PU4
I expect that it will be easy to submit my final project successfully with an autograder	PEOU1
I expect that when my code is run by the autograder, it will produce the same results as if I ran the same code on my local machine	PEOU2
I expect that using an autograder will not add a significant amount time spent on a final project	PEOU3
I expect that I will be able to get timely assistance when I run into technical problems with an autograder	PEOU4
I would choose an autograder for my final project instead of a human grader	BI2
I would prefer a human grader for my final project instead of an autograder	BI1

Table 1. Survey Questions

RESULTS

Factor Analysis

SPSS was used to perform factor analysis to verify the items loaded as predicted on the constructs generated fromTAM. This process was performed by selecting the elements for each factor, and ensuring that they loaded on a single factor.

Component Matrices					
Item	**Component**	**Item**	**Component**	**Item**	**Component**
PEOU1	0.771	**PU1**	0.793	**BI1**	-0.917
PEOU2	0.719	**PU2**	0.777	**BI2**	0.917
PEOU3	0.596	**PU3**	0.570		
PEOU4	0.565	**PU4**	0.790		

Table 2. Factor Loadings

The Behavioral Intention (BI) factors loaded strongly, with the positive and negativly wording questions loading appropriately. The Perceived Ease of Use (PEOU) factors all loaded strongly as well as the perceived usefulness factor (PU) all loaded strongly on the expected factor.

Hypothesis Testing

A Spearman's rho rank order correlation was used to determine the correlation between the survey items and the behavioral intention scores. As the behavioral intention items were worded both positively and negatively, and loaded strongly and similarly in the factor analysis, the positively worded version was used for hypothesis testing. This approach also allows the results to be explored from a practical perspective, as the specific correlations of each item, even though they were derived based on the underlying TAM model, can be explored and used to make practical suggestions.

(Partially Supported) *H1: Perceived Ease of Use will be positively correlated with behavioral intention for students using the autograder for large projects.* Items PEOU1 and PEOU3 both showed a positive correlation with the BI item at a significance level of .05. The other two items, PEOU2 and PEOU4 did not meet statistical significance

	PEOU1	**PEOU2**	**PEOU3**	**PEOU4**
BI2	.248*	0.121	.223*	0.137

* significant at .05 level

** significant at .01 level

(Supported) *H2: Perceived Usefulness will be positively correlated with behavioral intention for students using the autograder for large projects.* All Perceived Usefuleness Items (PU1, PU2, PU3 and PU4) showed a positive a significant correlation significant at the .01 level.

	PU1	**PU2**	**PU3**	**PU4**
BI2	.362**	.414**	.494**	.261**

* significant at .05 level
** significant at .01 level

(Supported)*H3: Perceived Ease of Use will not be significantly correlated with behavioral intention for course staff using the autograder for large projects*. None of the correlations for the staff were found to be significant for the perceived ease of use factors PEOU1, PEOU2, PEOU3 and PEOU4. This supports the findings of Gibson et al. (2008) that Perceived Ease of Use added little to their models predictive power.

	PEOU1	PEOU2	PEOU3	PEOU4
BI2	0.154	0.244	0.294	0.378

* significant at .05 level
** significant at .01 level

(Partially Supported) *H4: Perceived Usefulness will be positively correlated with behavioral intention for course staff using the autograder for large projects*. Factors PU1, PU2 and PU3 showed strong positive correlations, with PU1 and PU3 significant at the .05 level and PU2 significant at the .01 level.

	PU1	PU2	PU3	PU4
BI2	.507*	.655**	.424*	0.367

* significant at .05 level
** significant at .01 level

DISCUSSION

Based on our findings, it is obvious that student's primary concerns in the context of Perceived Ease of Use are primarily related to time and effort, as items PEOU1 and PEOU3 were both found to be positively and significantly correlated with the Behavioral Intention item, while the other two items did not show the same significance, and addressed other ease of use concepts.

Question	Factor
I expect that it will be easy to submit my final project successfully with an autograder	PEOU1
I expect that using an autograder will not add a significant amount time spent on a final project	PEOU3

Table 3. Significant PEOU Items for Students

In terms of Perceived Usefulness, all four items were found to be significantly and positively correlated with Behavioral Intention. Three of these items specifically focused on the usefulness aspect of the autograder to provide feedback, while the fourth focused on how the autograding experience was related to the final grade. These findings indicate that students are primarily concerned with the ability to get timely and valuable feedback from an autograder, something that may not be as available with the alternative of a human grader.

With respect to course staff and Perceived Ease of Use, our findings replicate the study of Gibson et al. (2008), finding that the relationships were not significant, and thus unlikely to add much in terms of predictive and explanatory power. Course staff, much like students, are however, concerned with many aspects of the systems Perceived Usefulness as three of the items had significant and positive correlations with the behavioral intention factor. Upon inspection of the items, the only item that was not supported had to do with the ability to receive easy to understand feedback. As course staff is more interested in providing students with feedback instead of understanding it themselves, we believe that this may be responsible for the lack of significance for this item.

LIMITATIONS

This study explores individual items developed based around the constructs from the Technology Aceptance Model (TAM) of Perceived Ease of Use (PEOU), Perceived Usefulness (PU) and Behavioral Intention (BI) in the specific context of autograders. The items were developed with the goal of being relevant and applicable to practitioners making decisions about implementing CBA. Specifically, to assess the importance of the individual items, while also generating a study based on the theory of the Technology Acceptance Model (TAM), Spearman Rho rank order correlations were performed using each item against a single measure of behavioral intention. These results are easier to interpret from a practitioner perspective, but may lack validity compared to some of the more rigorous analysis techniques.

CONCLUSION

This study extended past research applying the Technology Acceptance Model (TAM) to autograders, a form of Computer Based Assessment (CBA) by exploring it in the context of large projects, from both the perspectives of students and course staff. Our findings support the application of the Technology Acceptance Model in this domain, and provide partial replication of past studies in this area, specifically, the lack of support for the importance of Perceived Ease of Use (PEOU) for course staff as found in (Gibson et al. 2008). Future research on this topic could utilize more advanced statistical methods, and expanded versions of the technology acceptance model (TAM).

ACKNOWLEDGMENTS

We want to thank the project mentor and the reviewers in the OMSCS-CS6460-Ed Tech class at Georgia Tech.

REFERENCES

1. Allen, I. E., and Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the United States*. ERIC.
2. Atif, A., Richards, D., Busch, P., and Bilgin, A. (2015). "Assuring Graduate Competency: A Technology Acceptance Model for Course Guide Tools," *Journal of Computing in Higher Education* (27:2), pp. 94-113.
3. Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P. A., Fiset, M., and Huang, B. (2004). "How Does Distance Education Compare with Classroom Instruction? A Meta-Analysis of the Empirical Literature," *Review of educational research* (74:3), pp. 379-439.
4. Borokhovski, E., Bernard, R. M., Tamim, R. M., Schmid, R. F., and Sokolovskaya, A. (2016). "Technology-Supported Student Interaction in Post-Secondary Education: A Meta-Analysis of Designed Versus Contextual Treatments," *Computers & Education* (96), pp. 15-28.
5. Cheung, R., and Vogel, D. (2013). "Predicting User Acceptance of Collaborative Technologies: An Extension of the Technology Acceptance Model for E-Learning," *Computers & Education* (63), pp. 160-175.
6. Dasarathy, B., Sullivan, K., Schmidt, D. C., Fisher, D. H., and Porter, A. (2014). "The Past, Present, and Future of Moocs and Their Relevance to Software Engineering," *Proceedings of the on Future of Software Engineering*: ACM, pp. 212-224.
7. Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), pp. 319-340.
8. Edmunds, R., Thorpe, M., and Conole, G. (2012). "Student Attitudes Towards and Use of Ict in Course Study, Work and Social Activity: A Technology Acceptance Model Approach," *British journal of educational technology* (43:1), pp. 71-84.
9. Gibson, S. G., Harris, M. L., and Colaric, S. M. (2008). "Technology Acceptance in an Academic Context: Faculty Acceptance of Online Education," *Journal of Education for Business* (83:6), pp. 355-359.
10. Green, K. C., and Gilbert, S. W. (1995). "Great Expectations: Content, Communications, Productivity, and the Role of Information Technology in Higher Education," *Change: The magazine of higher learning* (27:2), pp. 8-18.
11. Ihantola, P., Ahoniemi, T., Karavirta, V., and Seppälä, O. (2010). "Review of Recent Systems for Automatic Assessment of Programming Assignments," *Proceedings of the 10th Koli Calling International Conference on Computing Education Research*: ACM, pp. 86-93.
12. Latchman, H., Salzmann, C., Gillet, D., and Bouzekri, H. (1999). "Information Technology Enhanced Learning in Distance and Conventional Education," *IEEE Transactions on Education* (42:4), pp. 247-254.
13. Means, B., Toyama, Y., Murphy, R., and Baki, M. (2013). "The Effectiveness of Online and Blended Learning: A Meta-Analysis of the Empirical Literature," *Teachers College Record* (115:3), pp. 1-47.
14. Park, S. Y. (2009). "An Analysis of the Technology Acceptance Model in Understanding University Students' Behavioral Intention to Use E-Learning," *Educational technology & society* (12:3), pp. 150-162.
15. Persico, D., Manca, S., and Pozzi, F. (2014). "Adapting the Technology Acceptance Model to Evaluate the Innovative Potential of E-Learning Systems," *Computers in Human Behavior* (30), pp. 614-622.
16. Schepers, J., and Wetzels, M. (2007). "A Meta-Analysis of the Technology Acceptance Model: Investigating Subjective Norm and Moderation Effects," *Information & management* (44:1), pp. 90-103.
17. Terzis, V., and Economides, A. A. (2011). "The Acceptance and Use of Computer Based Assessment," *Computers & Education* (56:4), pp. 1032-1044.
18. Venkatesh, V., and Davis, F. D. (2000). "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management science* (46:2), pp. 186-204.
19. Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). "User Acceptance of Information Technology: Toward a Unified View," *MIS quarterly*), pp. 425-478.
20. Wilcox, C. (2015). "The Role of Automation in Undergraduate Computer Science Education," *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*: ACM, pp. 90-95.