

# Exploring Digital Badges in University Courses: Relationships between Quantity, Engagement, and Performance

Joseph R. Fanfarelli & Rudy McDaniel  
*University of Central Florida*

## Abstract

Digital badging research is gaining momentum as instructors and administrators consider new models for assessing learning in nontraditional contexts (e.g., informal science learning programs, flexible online courses, adaptive learning systems). While many studies are examining the effectiveness of digital badges for pedagogical functions, such as motivating students, few attempts have been made to identify relationships between the number of badges earned by a student and badging effectiveness. The exploratory correlational study presented within this article addresses this gap by examining these relationships, relating number of badges earned in a pilot course to performance and engagement-related metrics. The results are further categorized by demographic groups to identify starting points for future research. Several relationships were identified, providing initial evidence of the importance of studying number of badges earned and how that number impacts effectiveness. The evidence suggests value in conducting large sample empirical research on the presented factors.

*Keywords:* badges, badging, achievements, metrics, engagement, individual differences, correlational

Fanfarelli Joseph R. & McDaniel, Rudy (2017). Exploring digital badges in university courses: relationships between quantity, engagement, and performance. *Online Learning 21* (2) doi: 10.24059/olj.v21i2.1007

## Introduction

Digital badges are digital images obtained through the completion of some pre-specified goal that are annotated with metadata (Grant, 2016). They are the digital correlates to merit badges and ribbons found in scouting and military organizations and used to reward achievement in specific activities. In this way, they function as “visual symbols of credentials” (Hickey et al., 2014) and as networked technologies with “the potential to both recognize and connect learning across contexts” (Davis & Singh, 2015, p. 72). As graphical indicators of learner activity, digital badges can function in online environments to validate performance in accomplishment or skill or provide evidence of quality of work or learner interest (HASTAC, n.d.). A benefit of digital

badges is that they are familiar technologies already used by many learners in other activities. For example, they have been used in entertainment applications, such as video games, for many years. In these environments, they are used to motivate players and credential their performance as “achievements” or “trophies” (Jakobssen, 2011; Hamari & Eranti, 2011) that can be viewed by other players.

Although digital badges are gaining momentum in online learning and instructors seem interested in using them, best practices for their effective use in pedagogy are still being investigated (Friedman, 2014). For example, in the domain of online learning, a number of companies and nonprofit organizations are exploring the use of digital badging in virtual coursework (e.g., Moskal, Thompson, & Futch, 2015). In recent years, digital badges have proven a popular subject in recent educational editorials and technology-focused news articles in publications such as *The Chronicle of Higher Education* (Young, 2012), *U.S. News & World Report* (Friedman, 2014), and *The New York Times* (Carey, 2012), all of which investigate the use and impact of digital badges in emerging pedagogical practices like micro-credentialing and information learning. Such work emerges from an interest in exploring alternate ideas to traditional educational assessment credentials such as course credits and college diplomas and in special educational programs like afterschool learning (Davis & Singh, 2015).

In the commercial sector, companies such as Credly (Credly, n.d.) now offer web-based systems for credentialing and validating online learning using a badge-based system. In the nonprofit space, Mozilla’s Open Badges initiative (Mozilla, n.d.), supported by the MacArthur Foundation, provides a suite of free software enabling participants to earn, deploy, and display badges in various learning scenarios. Earned badges can then be displayed in the Mozilla “backpack,” a website learners can use to share their earned credentials with peers, instructors, or even potential employers (Friedman, 2014). The Mozilla Foundation in particular is a strong advocate for digital badging, garnering support from leading business and learning organizations including Purdue University, Carnegie Mellon University, the University of California, the Smithsonian, Intel, and Disney-Pixar (Carey, 2012). Some institutions, such as George Washington University, even offer online MOOCs which allow members of the general public to enroll in courses and earn badges (Friedman, 2014). Such collaborations are invested in the idea that digital badges can positively transform the way we assess and credential learning, particularly in emerging technological landscapes.

In the academic literature, existing scholarly literature discusses digital badging from a number of perspectives. In terms of specific functionality, digital badges have shown potential as digital credentials (Gibson et al., 2013), mechanisms for engagement, motivation, and goal setting (Abramovich, Schunn, & Higashi, 2013), and as part of a larger gamified system (Su & Cheng, 2015), among other purposes (e.g., Antin & Churchill, 2011; Ulrich & Karvonen, 2011). Comprehensive reviews exploring the psychological dimensions of badging and best-practices for their design have also been conducted (McDaniel & Fanfarelli, 2016). Case studies about digital badges are also emerging; in one recent study, authors noted as both positive and negative reactions by students in college-level English composition courses (Reid & Paster, 2016). Students who enjoyed the badges mentioned that the badges helped them understand the directions in which they needed to work on revising their papers and praised the alignment between badge topics and course content. On the other hand, students who were frustrated by the badges described them as childish, difficult, and time-consuming. One respondent, for example,

noted that “this [the badging system] seems like a great idea for high school freshmen, not college level” (Reid & Paster, 2016, p. 198).

Nonetheless, despite the often-mixed reactions from both students and instructors, digital badges remain a popular method of augmenting coursework and directing student activities in specific ways. To investigate how to design them more effectively, researchers have begun identifying strategies for optimizing badging design both broadly (Hickey et al., 2014) and in specific contexts such as social media (Fanfarelli, Vie, & McDaniel, 2015). The Badges Design Principles Documentation Project (Hickey et al., 2014), for example, collects and captures the design principles used by winners of the 2012 Badges for Lifelong Learning Initiative, a competition sponsored by Mozilla and the MacArthur Foundation. The report also includes two case studies of badge systems used in education. One system was used for technology workplace skill development in a project called MOUSE Wins! and the other was used to train history teachers in a project called Who Built America (Hickey et al., 2014). This work resulted in a number of specific design principles for recognizing, assessing, and motivating learning. In addition, the report also suggested a number of principles for studying learning in digital badging systems, calling for researchers to think about the impact and use of badges in various ways as well as asking them to gather evidence of their successes and failures using a variety of research methods, one of which includes summative studies of digital badges at work within particular learning ecosystems.

While this type of summative research has been conducted to make badges more effective, it has also raised questions about whether or not badges are effective at all; for example, sometimes badges fail to enhance the constructs they seek to support (Fanfarelli, 2014). In examining the design aspects of badges, research hopes to uncover a series of strategies for creating badges and badging systems that are more consistently effective. This current manuscript seeks to fill a gap in the existing scholarly literature by providing quantitative data concerning the effectiveness of digital badges in online learning. The term *effectiveness* may be defined in a number of ways, and this article considers this term in a broad sense. Specifically, this study defines effectiveness as the degree to which digital badges produce a desirable result (e.g., higher levels of interest, enjoyment, or grades).

One way in which effectiveness may be studied regards the number of badges a student earns, an approach taken by Abramovich, Schunn, and Higashi (2013). They sought to gain a better understanding of how many badges were needed to form an effective badging system by examining the interaction between number of badges earned and type of motivation. Their middle school sample was split into two groups, high-performing and low-performing students, based on individual performance on a math pre-test. The researchers found that for low-performing students, a greater number of earned badges was associated with reduced performance avoidance orientation (i.e., students who prefer to avoid the appearance of underperforming, in contrast to those actively seeking to improve performance). For high-performing students, a greater number of earned badges was associated with increased expectancy to do well in math.

While correlational in nature and unable to support causal conclusions, this research raises questions about the influence of number of badges earned on badging effectiveness. Number of badges is likely to have at least a minor effect—the fact that badges can be effective, at all, suggests that the required number of badges is greater than zero; at least one badge is

required to show that badging can be effective. However, it is also likely that a single badge is insufficient to maximize effectiveness. On the other end of the spectrum, there will likely be a plateauing of effectiveness; a difference is unlikely to exist between 100,000 and 1,000,000 badges earned, and too many badges may even cheapen the experience, thereby reducing effectiveness. These extreme hypothetical values provide little guidance for the designer who is attempting to create a badging system that contains a realistic number of badges. Correlational studies, such as Abramovich et al.'s, lend insight into this problem, but few works have expanded upon this approach. The present article seeks to do just that. We present findings from a web-based university course to further examine the relationships between number of badges earned and other constructs. This work extends that done by Abramovich et al. by first examining correlations between number of badges earned and the constructs of the National Survey for Student Engagement (NSSE, 2015). These constructs include Student Satisfaction, Learning Strategies, Effective Teaching Practices, Higher Order Learning, Student Faculty Interaction, Reflective and Integrative Learning, and Collaborative Learning. The NSSE will be described in greater depth in the methodology section.

After identifying these relationships, we examined the correlations at a more specific level. The effects of individual differences on badging effectiveness are not well studied. To contribute to this area of research, we categorized data by gender, perceived importance of badging systems, and prior frequency of interaction with badging systems to begin to identify the role these constructs play in badging effectiveness.

Finally, we examined the potential of badges to serve as assessment tools. Badges in these roles benefit those looking to implement badges as credentials (Gibson et al., 2013). Credentialing badges can be used to provide a finer level of granularity to assessment (e.g., a badge that was awarded for excellence in developing an artificially intelligent pathfinding algorithm tells more about a student's specific skillset than an A in their Game Programming course).

Badging also benefits those looking for alternatives to traditional formal exams when, for example, test anxiety is a concern. Test anxiety can be defined as the set of negative cognitive, behavioral, and physical responses that occur when an individual is concerned about the possibility of failure on an examination (Zeidner, 1998), and has been deemed responsible for a range of negative effects on academic performance measures (Cassady & Johnson, 2002). This poses a problem. If test anxiety is a cause of reduced performance, it is reasonable to conclude that tests and assessments are not measuring an individual's actual skills or abilities; instead, they are only measuring the individual's capabilities during a moment of weakness, leading to inaccurate assessment. Nevertheless, assessment remains an important process that reflects the instructor's effectiveness as well as the student's current level of competency (Guskey, 2003). To maintain the benefits of assessment while reducing the potential for test anxiety, badging can serve as a form of covert assessment. As students demonstrate their competencies throughout the course of a semester, badges can be awarded and tied to the assignments or activities in which students demonstrated the competency. In this manner, the badging process becomes a type of informal, continuous, and traceable assessment, providing similar benefits to traditional assessment, without the stress that is placed on students during formal examination.

However, the 100-point grading scale that is used to assign value to performance in traditional assessment is a time-tested tool, and may prove difficult to abandon for many educators. Regardless, a thorough search of the literature revealed examples of badges being used in this manner (Rapti, 2013; Terrell, 2014), but no research was found empirically examining the potential for badges to be used as grade predictors. Our discussion contributes to an understanding of the potential for the number of badges earned metric to be used as a predictor of final grade. Our research questions are as follows:

*Research Question 1:* To what extent is number of badges earned correlated to final grade, intrinsic motivation, engagement, and the constructs measured by the National Survey for Student Engagement (student satisfaction, learning strategies, effective teaching practices, higher order learning, student faculty interaction, reflective & integrative learning, collaborative learning)?

*Research Question 2:* How do gender, perceived importance of badging systems, and prior frequency of interaction with badging systems impact the correlations between number of badges earned and badging effectiveness?

*Research Question 3:* Can number of earned badges be used to predict a numerical final grade on a 100-point scale?

These research questions will be examined in the context of four web-based courses in the Digital Media department at the University of Central Florida. We hope the results of this study will guide future large scale research toward appropriate metrics.

## **Method**

### **Participants**

This study was included as part of a larger experiment comparing the effectiveness of badged and non-badged interventions. The complete dataset can be found in McDaniel and Fanfarelli (2015). While the larger experiment included forty-four participants, twenty-one (13 male and 8 female) were included in this study (i.e., the participants from the “badged” condition). All participants were 18 years or older, ranging from first semester freshman to final semester seniors at the University of Central Florida. Participants were recruited from two sections each of two different web-based elective courses (Web Design and Graphic Design), all taught by the same instructor.

### **Course Design**

Courses were open to all students enrolled in the university’s School of Visual Arts and Design. Both courses were created to cater to both novice and expert users of either Adobe Photoshop (Graphic Design) or Adobe Dreamweaver (Web Design), and prepared students to take the Adobe Certified Expert certification exam. Courses were designed to place equal emphasis on in-course examinations and project-based assignments, each receiving a 50% grade weighting. Each course was offered fully online through the Canvas course management system, and included 10 assignments, 11 quizzes, three exams, and a cumulative final exam. Quizzes and exams utilized a multiple-choice format, and were instantly graded by the system, providing instantaneous feedback.

## Badge Design

Figure 1 contains examples of badge designs used in this study. A complete list of badges and their explanations can be located within the online repository included in McDaniel and Fanfarelli (2015).



Figure 1. Three of the badges used in this study.

A total of 22 badges were designed for inclusion in the experimental courses (18 of these were included in Graphic Design and 19 were included in Web Design). However, only 18 were awarded to at least one student (14 in graphic design and 15 in web design). It was originally anticipated that a clear “top three” grades would emerge on some assignments, but this did not manifest in practice, with multiple students consistently securing each of the top three positions. For this reason, badges for attaining the highest, second highest, and third highest assignment scores were never awarded. It was also anticipated that some students would obtain a perfect score on all quizzes – a condition that was also never met, removing another badge from the set.

Badges were designed to be unexpected and skill-based, in contrast to being expected and awarded for the mere completion of a mandatory task, to avoid the de-motivational effects that are occasionally associated with rewards (Cameron & Pierce, 1994) and to increase the likelihood of motivational benefit (Abramovich, Schunn, & Higashi, 2013). The term *unexpected badges* refers to badge systems that hide the badge and its criteria until the badge is earned. In other words, users are unable to forecast when they will earn a badge, and may be completely unaware of its existence. Thus, when it is earned, it is an unexpected surprise. *Skill-based badges* refers to those badges that are earned through the demonstration of some skill or ability. In other words, they are not handed out for mere participation. These badges tend to be rarer and have higher associated value than participatory badges.

The system utilized one badge that was not skill-based. It was an introductory badge given to all students upon the first day of the course. This badge was included to notify students of the badge system’s existence. This was necessary because badges were designed to be unexpected. Without the introductory badge, students would have had no indication that they could even earn badges within the course.

Upon earning a badge in the course, students were sent an e-mail that notified them of their badge. The badge could then be viewed alongside their other earned badges within the learning management system that was used for their normal course-related activities. For each badge, the badge’s graphic was displayed, along with a description of the badge, and a message

describing how it was earned. Participants were unable to see which badges their classmates had earned.

Badges were awarded in two different ways. Those related to the grade on a single quiz or exam were assessed by the system and automatically awarded upon completion of the assessment. All other badges were awarded by the instructor, who awarded them promptly (e.g., the pre-emptive strike badge was given for completing an exam at least two days before it was due and receiving a 90% or higher. This badge was awarded at midnight on the day before the exam was due).

## **Materials**

Participants completed the following questionnaires during the course of the experiment.

**Demographics questionnaire.** The Demographics Questionnaire assessed how often participants interacted with badges prior to experimentation (frequency of interaction), how important participants considered badging to their gaming experience (importance), and information regarding gameplay habits, experience, and genre preferences. Importance was assessed in relation to gaming because gaming was believed to be the avenue that was most likely to have facilitated interactions between participants and badges, due to the frequent use of badges in games (Abramovich, Schunn, & Higashi, 2013).

**National Survey for Student Engagement (NSSE).** The NSSE assesses the extent to which students engage in effective educational practices. A modified version of the NSSE (NSSE, 2014) was included to assess various factors related to student engagement. Factors of specific interest in this study included collaborative learning, reflective and integrative learning, student faculty interaction, higher order learning, effective teaching practices, learning strategies, and student satisfaction. The NSSE was selected for two reasons. First, it efficiently measured a range of variables of interest within a single questionnaire. Second, it is used extensively in educational benchmarking, making it a tool familiar to many educators. Modifications were made to better suit the structure of the course. For example, one of the questions asked if the student gave a course presentation. This was not an option in this course, and all students would have responded with “Never.” This question, and others that were not applicable, were thus omitted. See Appendix A: Modified NSSE Questionnaire to view the version used for this study.

**Interest/Enjoyment Subscale of the Intrinsic Motivation Inventory (IMI) (Ryan, 1982).** The IMI measured intrinsic motivation. Intrinsic motivation defines a person’s internal desire to complete a task (e.g., sense of fulfillment, personal pride, internal joy, etc.). This is in contrast to extrinsic motivation, which defines a person’s desire to complete a task to obtain some external reward (e.g., money). By measuring intrinsic motivation, we were able to better understand how badges did or did not motivate students.

**Engagement Measure (Jennett et al., 2008; Charlton & Danforth, 2005).** The Engagement Measure identified participants’ level of engagement during the course. Engagement is one possible product of motivation (Walker, Greene, & Mansell, 2006). After a learner is motivated, he/she can then engage in the task and willfully exert meaningful effort. Measuring engagement enables the differentiation between learners who are engaged, and those who are motivated but not engaged.

**Perceived Importance of Badging Systems.** We thought it would be interesting to identify whether badges differently affected those who had preconceived biases toward badges versus those who did not. To identify whether this bias existed, we asked, “If you play video games, how important are achievements / trophies / badges in improving your gaming experience?” and allowed users to answer on a five-point Likert scale, with an additional “I do not play video games” option. Video games were chosen for this question because we believed this was the context in which students were most likely to have positively encountered badges. By choosing the video game domain, we hoped we might avoid opening the question to a wide variety of badging experiences that may or may not match the implementation of badges used in the course.

**Frequency of Interaction with Badges.** In addition to bias, we also wanted to investigate whether experience with badges was a mediating factor on effectiveness. We asked the question “How often do you interact with games or software that use achievements or badges?” and provided the options, “Daily,” “Weekly,” “Monthly,” “Once every few months,” “Rarely,” and “Never.”

Final grade and number of badges earned were also collected for each participant. Number of badges earned was used to gain insight into the relationships between the tested factors and badge quantity. Strong or weak relationships between these two factors provided insight into the design of badging systems for assessment.

## **Procedure**

Participants completed their respective courses as they would any other web-based course. Questionnaires were made available three weeks before the end of the semester to allow adequate time for completion. All questionnaires were completed online via Qualtrics, a secure survey system. Students were informed that they could access the survey at their leisure during the allotted time, and were not required to notify the instructor or experimenter that the surveys were completed.

## **Data Analysis**

This study implemented two forms of data analysis. First, correlation analysis was conducted to obtain the Pearson correlation coefficient,  $r$ . The correlation coefficient measures the strength of a linear relationship between two variables and was important for this study’s goal of better understanding the relationships between number of badges earned and the studied variables. It also enabled the calculation of the coefficient of determination,  $R^2$ , to identify the amount of variance in final grade that was explained by the number of badges earned metric.

Regression analysis is used to estimate relationships between variables, and was conducted to test the predictive ability of a single variable (number of badges earned) on an outcome variable (final grade). This methodology was chosen because it yielded formulae that could be examined and compared to formulae derived from future studies exploring relationships between final grade and number of badges earned. Ultimately, it is hoped that the derived formulae will be useful in predicting how badge-earning behavior may influence a student’s final grade when badge systems are used that are similar in nature to the one used in this study. *Similar*, in the case of digital badging, is a term that lacks a concrete definition—a larger body of experimental research is necessary before we can clearly understand how different badging systems can be without producing variable results.



Significance was observed at  $p < 0.05$ . In line with Cohen's (1988) recommendations,  $r > 0.30$  was considered a medium correlation, and  $r > 0.50$  was considered large. While Cohen also specified that correlations should be considered small at  $r > 0.10$ , no small correlations were found to be significant in this study. Accordingly, the coefficients for these variables are listed, but not discussed or submitted to further analysis.

## Results

### Correlations with Number of Badges Earned

All correlational data is presented in Table 1.

Correlations with Number of Badges Earned (r)							
	Global	High $\pm$ Frequency	Low Frequency	High Importance	Low Importance	Male	Female
Final Grade	***0.69	**0.58	***0.95	0.50	N/A	***0.69	0.60
Intrinsic Motivation	0.23	0.24	0.15	-0.40	N/A	0.09	0.39
Engagement	0.18	0.15	0.09	-0.20	N/A	-0.05	0.50
Student Satisfaction	**0.47	0.27	0.67	0.27	N/A	0.16	**0.73
Learning Strategies	0.34	0.35	0.39	-0.03	N/A	0.36	0.46
Eff. Teaching Practices	0.05	0.11	0.12	-0.25	N/A	-0.03	0.20
Higher Order Learning	0.14	0.11	0.25	-0.42	N/A	0.06	0.38
Student Faculty Interaction	0.04	-0.22	0.28	-0.39	N/A	-0.10	0.38
Ref. & Integ. Learning	0.22	0.50	-0.09	-0.34	N/A	0.27	0.07
Collaborative Learning	0.12	0.02	0.25	0.04	N/A	0.26	0.19

\*\*Significant ( $p < 0.05$ ) \*\*\*Significant ( $p < 0.01$ )

$\pm$ High Frequency refers to students who used badges at least once per month, prior to entering the course

**Final Grade.** Significant positive correlations were observed between number of badges earned and final grade when examined across all participants ( $r(21) = 0.69$ ,  $p < 0.001$ ). See Figure 2.

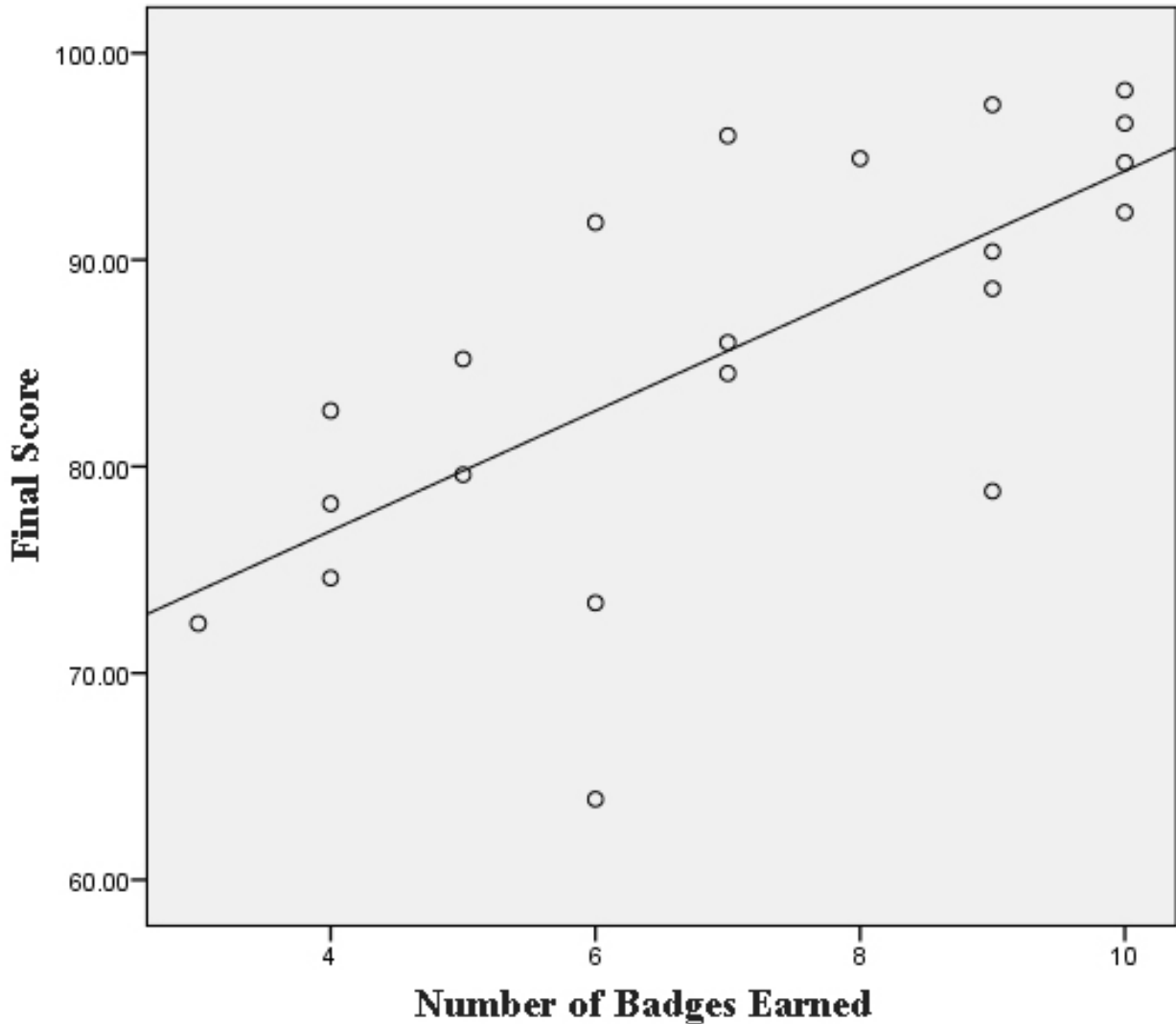


Figure 2. Correlation Between Number of Badges Earned and Final Score across All Participants.

Students who earned a greater number of badges were more likely to have a higher final grade than their low-earning counterparts. The strength of this relationship opens the possibility for calculating grades from earned badges and provides early evidence that badges may be useful as evaluation systems. To acquire more insight into the forces behind the correlation coefficient's value, the relationship was examined regarding participants' frequency of interaction, gender, and perceived badge importance.

No significant correlations were found for participants who considered badges to be important to the gaming experience, and there were not enough respondents to run analyses on participants who did not consider badges to be important to the gaming experience. However, some variations were discovered in other categories, providing guidance for when badge-based evaluation systems may be most useful. For example, male participants showed a strong significantly positive relationship between final grade and number of badges earned ( $r(13) = 0.69$ ,  $p = 0.01$ ), providing support for badge-based evaluation systems for male students. See Figure 3.

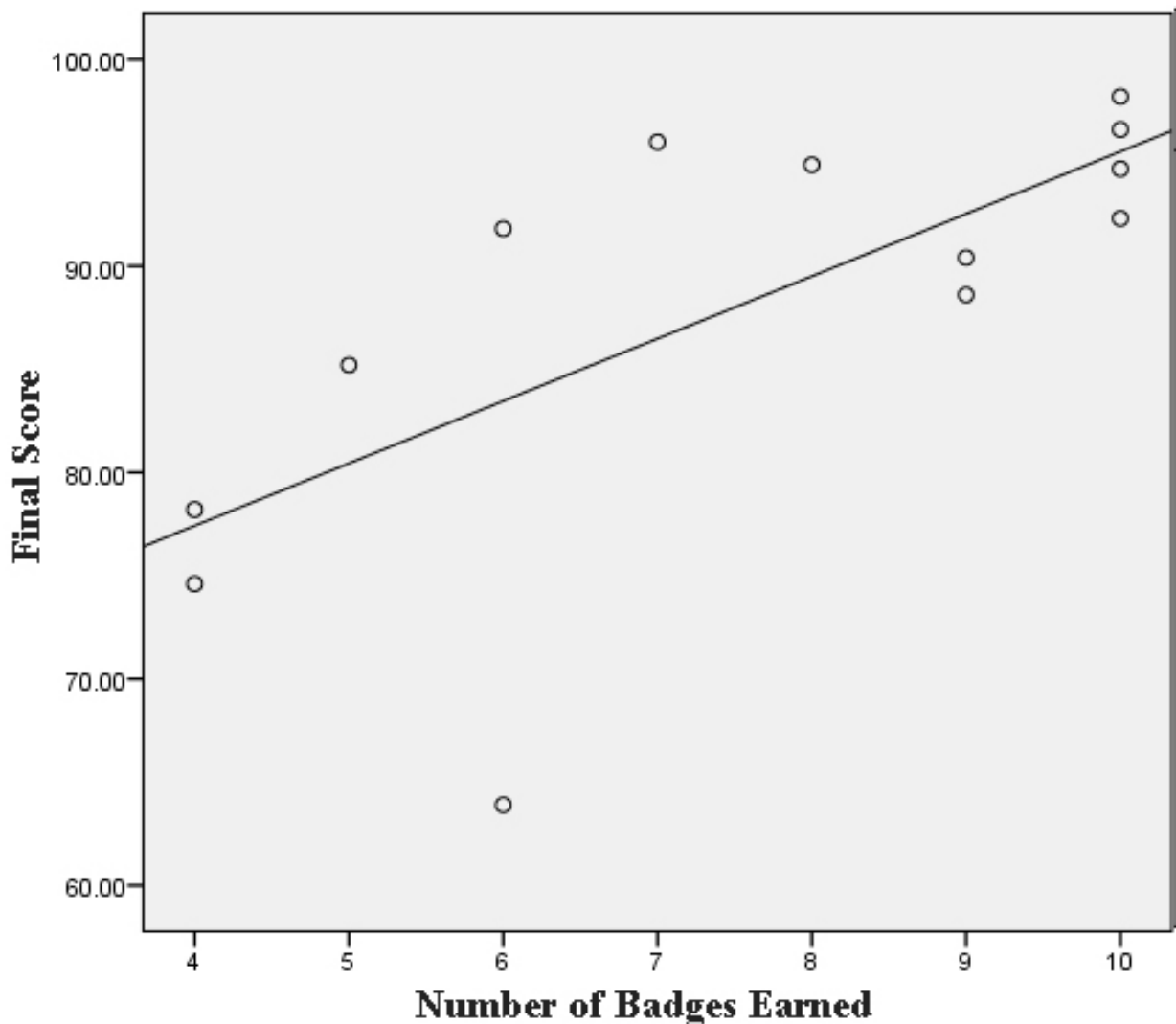


Figure 3. Correlation Between Number of Badges Earned and Final Score for Male Participants.

For female participants, however, the correlation between final grade and number of badges earned did not quite reach significance. While a relatively strong relationship was observed for female participants, there is not enough evidence to confidently state that this relationship was not due to chance. Accordingly, this study does not provide enough support for using badge-based evaluation systems in all-female courses. However, there is some hesitation in definitively stating that badge-based evaluation should not be used for females. The correlation coefficient was fairly large. Combined with this study's small sample size, a significant relationship may yet exist. Future studies should examine this relationship with a larger female sample size.

Frequency of interaction also revealed interesting results. See Figures 4 and 5.

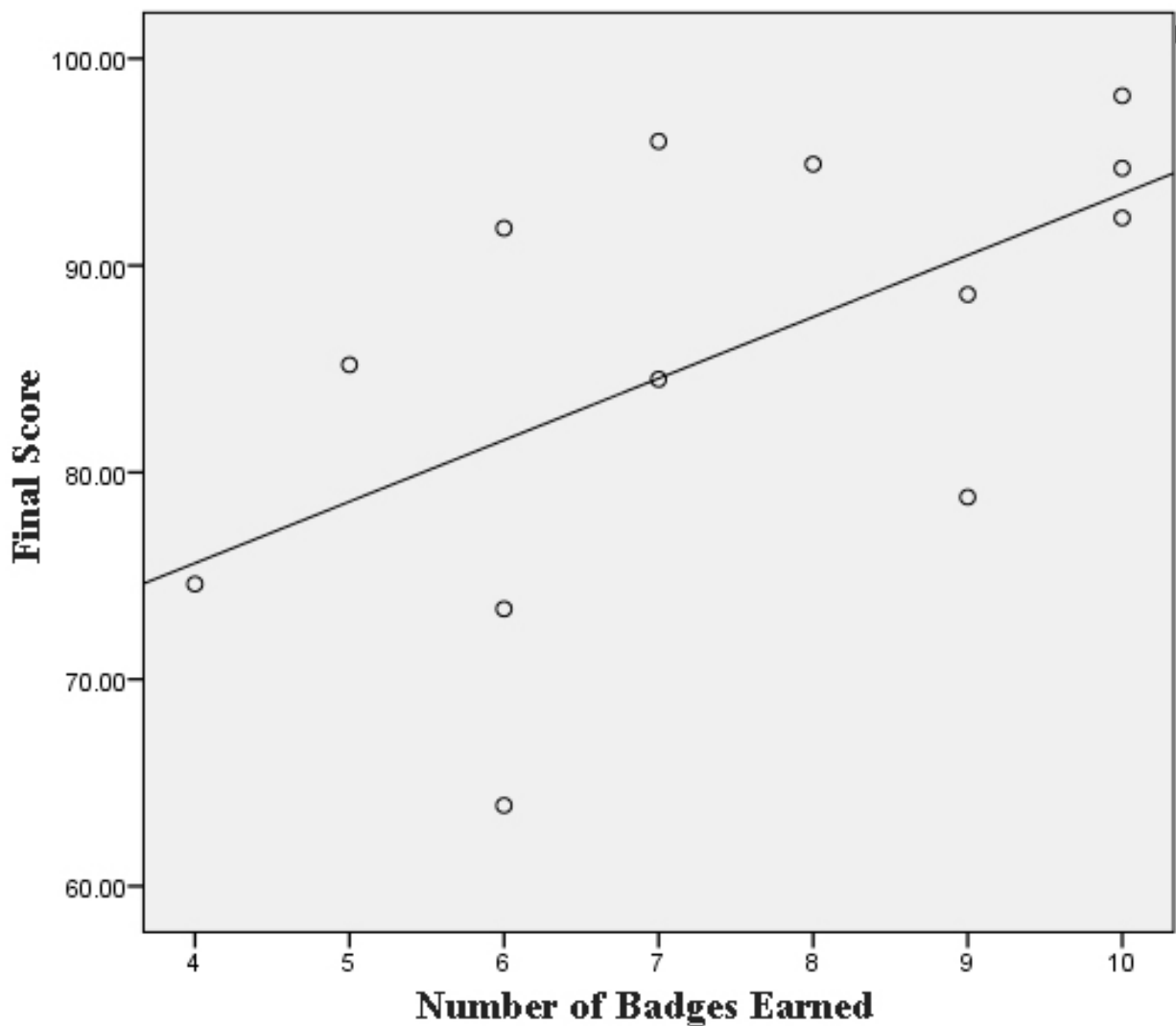
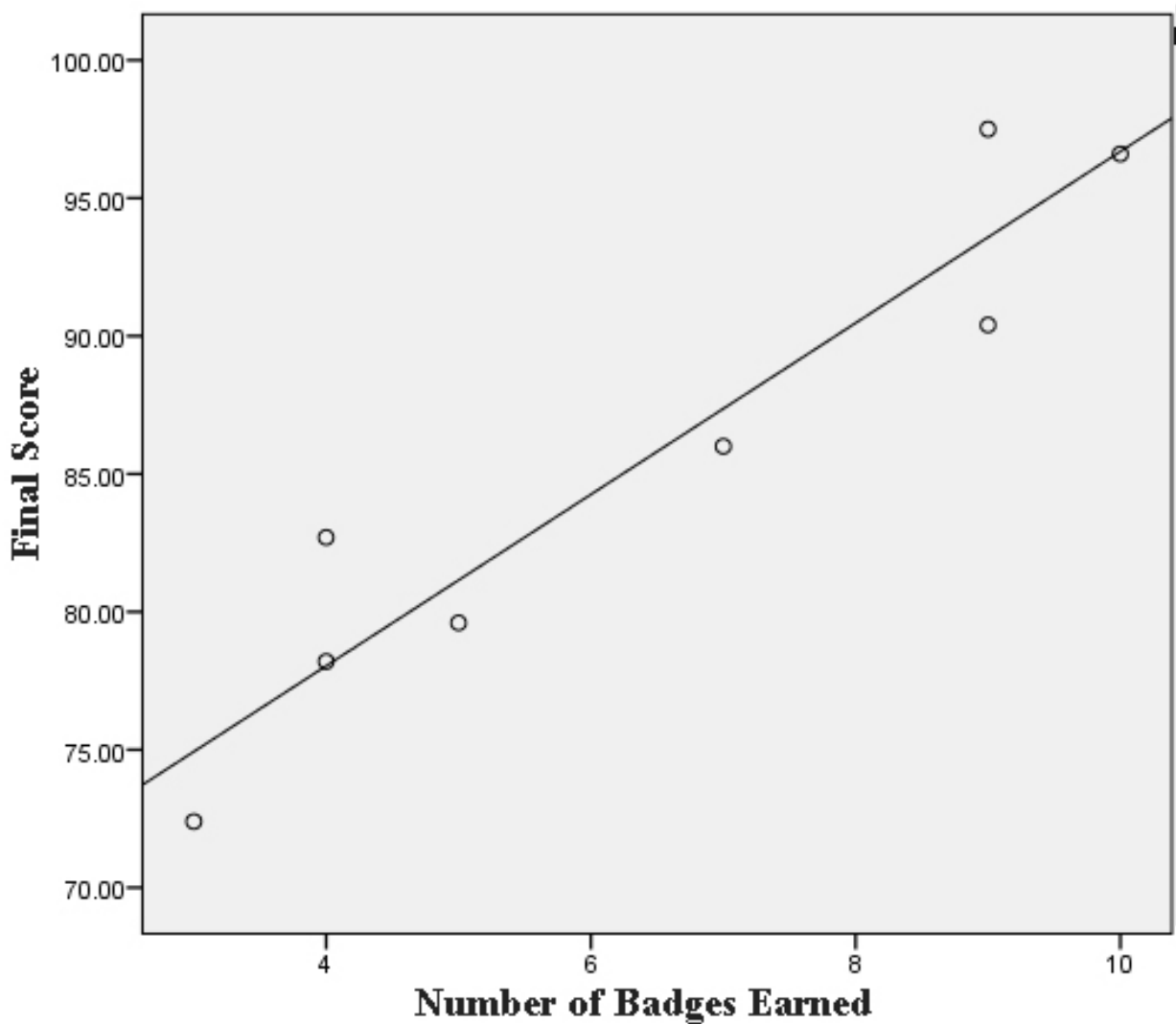


Figure 4. Correlation Between Number of Badges Earned and Final Score for High Frequency of Interaction Participants.



2

*Figure 5. Correlation Between Number of Badges Earned and Final Score for Low Frequency of Interaction Participants.*

While both groups displayed strong significantly positive relationships (high,  $r(13) = 0.58$ ,  $p = 0.04$ ; low,  $r(8) = 0.95$ ,  $p < 0.001$ ), participants who had less frequently interacted with badging systems prior to experimentation boasted an extremely strong relationship – approaching 1.0. A strong positive relationship was also found for high frequency of interaction participants, though to a lesser degree. A badge-based evaluation system may be effective for both groups, but an instructor who knows that their students have minimal experience with badging (e.g., those who are not technologically inclined) can implement such a system with a greater level of confidence.

**Student Satisfaction.** Significant correlations were also found between student satisfaction and number of badges earned when examined across all participants ( $r(21) = 0.47$ ,  $p = 0.03$ ), and female participants ( $r(8) = 0.73$ ,  $p = 0.04$ ). See Figures 6 and 7.

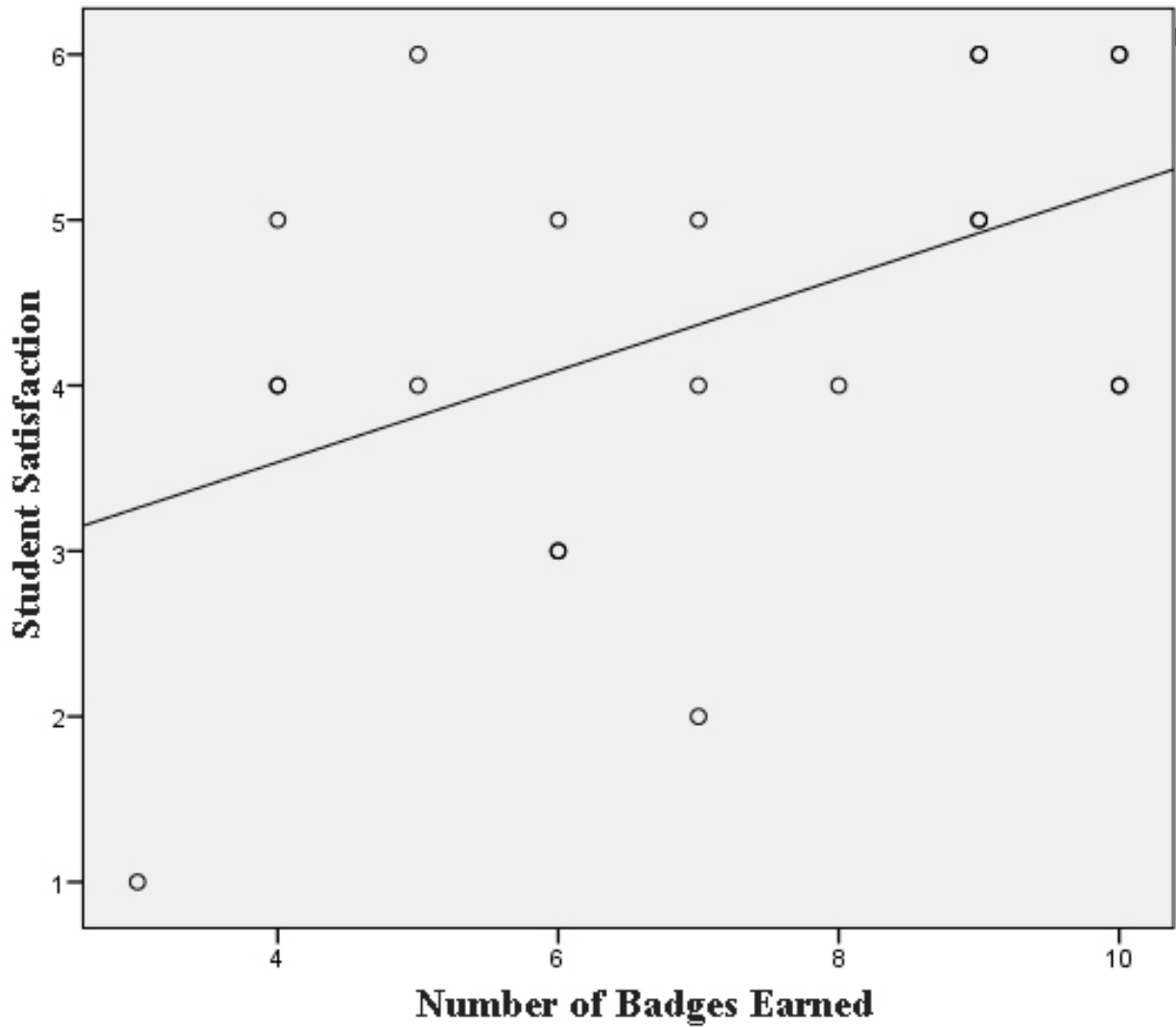


Figure 6. Correlation Between Number of Badges Earned and Student Satisfaction Across All Participants.

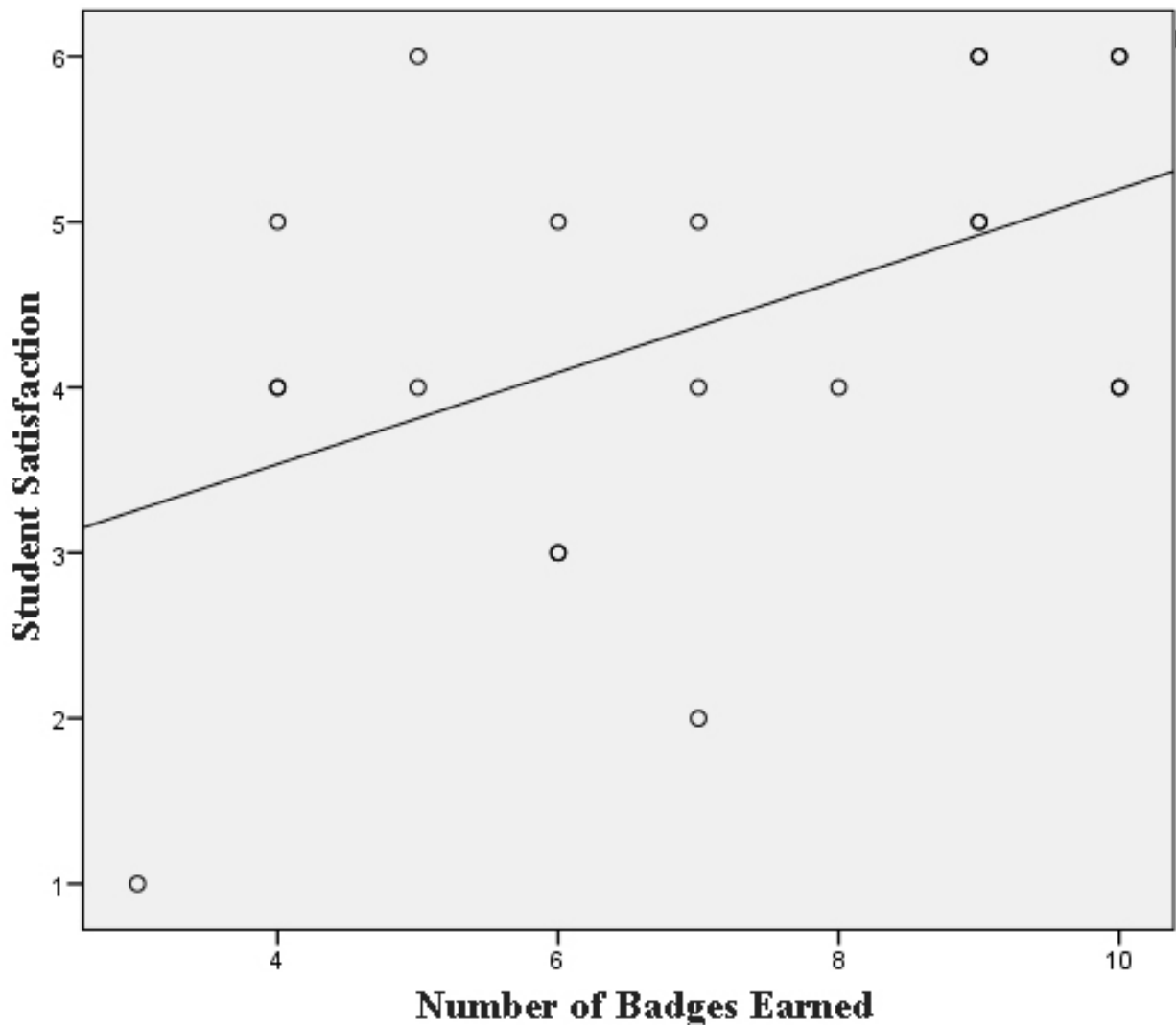


Figure 7. Correlation Between Number of Badges Earned and Student Satisfaction for Female Participants.

While a medium-strength positive relationship was found across all participants, female participants showed a much stronger relationship than other subgroups. In comparison, males showed a very small non-significant relationship at  $r = 0.16$ . These results fit within those obtained by McDaniel, Lindgren, and Friskics (2012), who observed that satisfaction with their badging system was higher for female participants than males. It seems that in addition to finding greater satisfaction in the existence of a badging system, female students may also derive greater satisfaction in acquiring more badges within the system. Future studies should implement controlled experimental methodologies to test this possibility. If the finding holds up through experimental rigor, badging systems geared toward a female population may be more effective when participants are able to earn a high number of badges.

Significant correlations were not found for intrinsic motivation, engagement, or any of the other NSSE variables.

### **Number of Badges Earned as a Predictor of Final Grade**

In response to the strong correlations identified in the previous section, number of badges earned was investigated as a potential predictor for final percentage grade, and was found to be a significant predictor of final score,  $\beta = 2.90$ ,  $p < 0.001$ . The following formula was obtained:

$$\text{Final grade} = 65.29 + 2.90(\text{NumberOfBadgesEarned})$$

Number of badges earned explained a significant proportion of the variance in final score,  $R^2 = 0.48$ ,  $F(1,19) = 17.55$ ,  $p < 0.001$ . In other words, 48% of the variance in final score was explained by number of badges earned. While significant, 48% is probably not an appropriate degree of accuracy for a grading environment in which grading accuracy is very important (e.g., most formal educational institutions). However, across all participants, the correlation was much lower than it was for a subgroup of the participants—those with low frequency of interaction scores. While badging may not be the optimal choice for alternative grading in formal educational institutions that cater to students from a variety of backgrounds, perhaps they could be useful in environments where students are less likely to have experience with badges.

To examine the possibility, regression analysis was conducted for low frequency of interaction participants (i.e., participants who used badges less than once per month prior to experimentation). Number of badges was again a significant predictor of final score,  $\beta = 3.11$ ,  $p < 0.001$ :

$$\text{Final grade} = 65.61 + 3.11(\text{NumberOfBadgesEarned}).$$

However, this time, number of badges earned explained a much larger proportion of variance in final score,  $R^2 = 0.90$ ,  $F(1,6) = 51.93$ ,  $p < 0.001$ . Number of badges earned explained 90% of the variance in final grade, making it much more reliable with this population. While 90% may still be too low for serious examination in a high-stakes environment, it could be sufficient for more informal learning environments, or for portions of less critical examination within a formal environment.

### **Discussion**

The results presented in this study are exploratory in nature and help define areas that show potential for large-scale empirical research, representing a step forward in the effort to use badges as assessment tools. The informal nature of badges and their inherent flexibility that allows them to be pre-defined to match desired course outcomes may make them a suitable tool for evaluation, especially when exact grades are unnecessary and covert assessment is desirable. However, educators should proceed with caution. Badging for assessment is unlikely to be a one-size-fits-all solution.

These results are to be a part of a larger effort to uncover how badges work in different contexts with different learners. For instance, the badging system used in this study was designed in accordance with recommendations from the literature—badges were skill-based, unexpected, and aligned to formal assessment. Badging systems that do not conform to this design philosophy, or apply the philosophy in different ways may obtain results that differ from those presented in this article. Future research will help to specify how differences in design affect outcomes.



There are a few particular limitations of this study—and of badging—which should be considered in future experiments. First, while this study sampled the target population in the target environment, it was constrained to a small sample size due to the nature of the courses tested. Future researchers should use the results of this study to guide the formation of larger scale studies on badges. This study provides evidence to support the idea that number of badges earned correlates with final grade. Now, rigorous, larger scale experiments on these metrics are both justifiable and necessary.

Additionally, the positive relationships discovered in this study should be further assessed to determine existence and direction of causation. If the act of earning more badges was responsible for increasing student satisfaction, badging system designers may want to consider creating systems with a greater number of badges. Of course, it is also possible that greater satisfaction caused students to earn more badges. This conclusion, while plausible, may be less likely, since students were unable to see which badges were available to be earned, hindering their ability to engage in directed badge-hunting. However, it is possible that satisfied students stumbled upon more badges as they put more effort into course assignments and examinations. Whichever is the case, a better understanding of the stimulating factors will be informative and may lead to new design guidelines and practices.

Another concern is that the exact predictive ability of number of badges earned may vary on a case-by-case basis, especially as course structure and specific badges vary. This experiment used a particular set of badges that utilized criteria designed by the authors. Criteria designed by others, or in different ways, may yield different results. Future research needs to examine the robustness of these results within other domains and with different sets of skill-based badges. We made an initial necessary effort on this front, but a full understanding of badges for this purpose will only be gained through a number of studies in different domains, and with different participants. We presented predictive formulae for the current study. It will be important for future researchers to present their own formulae to see if they globally exhibit similarities, or resemble each other within particular domains, or within badging systems of particular designs.

### References

- Abramovich, S., Schunn, C., and Higashi, R.M. (2013) Are badges useful in education? It depends upon the type of badge and expertise of the learner. *Educational Technology Research and Development*, 61, 217-232.
- Antin, J. and Churchill, E.F. (2011) Badges in social media: A social psychological perspective. In *Proceedings of CHI 2011*, Vancouver, BC, Canada.
- McDaniel, R., & Fanfarelli, J.R. (2016). Building Better Digital Badges Pairing Completion Logic With Psychological Factors. *Simulation & Gaming*, 47(1), 73-102.
- McDaniel, R. & Fanfarelli, J.R. (2015). A digital badging dataset focused on performance, engagement, and behavior-related variables from observations in web-based university courses [Special issue]. *British Journal of Educational Technology*, 46(5), 937-941.
- Cameron, J. and Pierce, D.W. (1994) Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, 64(3), 363-423.

- Carey, K. (2012). Show me your badge. *The New York Times*. Retrieved December 10, 2015, from <http://www.nytimes.com/2012/11/04/education/edlife/show-me-your-badge.html>
- Cassady, J.C. and Johnson, R.E. (2002) Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, 27, 270-295.
- Cohen, J. (1988) Statistical power analysis for the behavioral sciences (2<sup>nd</sup> ed.). New Jersey: Lawrence Erlbaum.
- Credly.com (n.d.). Credly: Achievements verified and shared. Retrieved December 10, 2015, from <https://credly.com/>.
- Davis, K., & Singh, S. (2015). Digital badges in afterschool learning: Documenting the perspectives and experiences of students and educators. *Computers & Education*, 88, 72-83.
- Fanfarelli, J. (2014). *The Effects of Narrative and Achievement on Learning in a 2D Platformer Video Game*. Doctoral Dissertation. University of Central Florida, Orlando, FL, USA.
- Fanfarelli, J.R., Vie, S., & McDaniel, R. (2015). Understanding digital badges through feedback, reward, and narrative: A multidisciplinary approach to building better badges in social environments. *Communication Design Quarterly*, 3(3), 56-60.
- Friedman, J. (2014). Online courses experiment with digital badges. *US News & World Report*. Retrieved December 10, 2015, from <http://www.usnews.com/education/online-education/articles/2014/12/10/online-courses-experiment-with-digital-badges>
- Gibson, D., Ostahewski, N., Flintoff, K., Grant, S., and Knight, E. (2013) Digital badges in education. *Education and Information Technologies*, 20(2), 403-410.
- Grant, S. L. (2016). History and context of open digital badges. In L. Y. Muilenburg & Z. L. Berge (Eds.), *Digital Badges in Education: Trends, Issues, and Cases* (pp. 3-11). New York: Routledge.
- Guskey, T.R. (2003) How classroom assessments improve learning. *Educational Leadership*, 60(5), 6-11.
- Hamari, J. & Eranti, V. (2011). Framework for designing and evaluating game achievements. *Proceedings of DiGRA 2011*, Hilversum, Netherlands.
- HASTAC: Humanities, arts, science, and technology alliance and collaboratory (n.d.) Digital badges. Retrieved from <http://www.hastac.org/digital-badges>.
- Hickey, D. T., Itow, R., Schenke, K., Tran, C., Otto, N., & Chow, C. (2014). Badges Design Principles Documentation Project. January Interim Report: Retrieved from Indiana University, the Badges Design Principles Documentation Project website: <http://iudpd.indiana.edu/JanuaryReport>
- Jakobsson, M. (2011). The achievement machine: understanding Xbox 360 achievements in gaming practices. *Game Studies*, 11(1), 1-22.
- McDaniel, R., Lindgren, R. & Friskics, J. (2012). Using badges for shaping interactions in online learning environments. *Proceedings of the IEEE International Professional Communication Conference*, 1-4. Orlando, FL, USA.

- Moskal, P. Thompson, K. & Futch, L. (2015). Enrollment, engagement, and satisfaction in the blendkit faculty development open, online course. *Online Learning Journal*, 19(4), 1-12.
- Mozilla (n.d.). OpenBadges. Retrieved from <http://openbadges.org/>.
- NSSE (2015) Survey Instrument. National Survey For Student Engagement. Retrieved from [http://nsse.iub.edu/html/survey\\_instruments.cfm](http://nsse.iub.edu/html/survey_instruments.cfm)
- Rapti, K. (2013) Implementing alternative assessment methods through gamification. *Proceedings of ICERI '13*, 255-262. Seville, Spain.
- Reid, A. J., & Paster, D. (2016). A case-study of digital badges in composition courses. In L. Y. Muilenburg & Z. L. Berge (Eds.), *Digital Badges in Education: Trends, Issues, and Cases* (pp. 189-202). New York: Routledge.
- Su, C-H. & Cheng, C-H. A mobile gamification learning system for improving he learning motivation and achievements. *Journal of Computer Assisted Learning*, 31(3), 268-286.
- Terrell, S.S. (2014) *The 30 Goals Challenge for Teachers: Small Steps to Transform Your Teaching*. London: Routledge.
- Ulrich, J. & Karvonen, M. (2011). Faculty instructional attitudes, interest, and intention: Predictors of web 2.0 use in online courses. *The Internet and Higher Education*, 14(4), 207-216.
- Young, J.R. (2012). 'Badges' earned online pose challenge to traditional college diplomas. *The Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/Badges-Earned-Online-Pose/130241/>.
- Zeidner, M. (1998) *Test Anxiety: The State of the Art*. New York: Plenum Press.

### Appendix A: Modified NSSE Questionnaire

This questionnaire was modified from the original NSSE, which can be found at the following location: [http://nsse.indiana.edu/html/survey\\_instruments.cfm?siFlag=yes&sy=2014](http://nsse.indiana.edu/html/survey_instruments.cfm?siFlag=yes&sy=2014)

#### 1. During the current semester, about how often have you done the following?

a. Asked questions or contributed to course discussions in other ways

Very often                  Often                  Sometimes                  Never

b. Prepared two or more drafts of an assignment before turning it in

Very often                  Often                  Sometimes                  Never

c. Attended an art exhibit, play or other arts performance (dance, music, etc.)

Very often                  Often                  Sometimes                  Never

d. Asked another student to help you understand course material

Very often                  Often                  Sometimes                  Never

e. Explained course material to one or more students

Very often                  Often                  Sometimes                  Never

f. Prepared for exams by discussing or working through course material with other students

Very often                  Often                  Sometimes                  Never

#### 2. During the current semester, about how often have you done the following?

a. Combined ideas from different courses when completing assignments

Very often                  Often                  Sometimes                  Never

b. Learned something that changed the way you understand an issue or concept

Very often                  Often                  Sometimes                  Never

c. Connected ideas from your courses to your prior experiences and knowledge

Very often                  Often                  Sometimes                  Never

#### 3. During the current semester, about how often have you done the following?

a. Talked about career plans with a faculty member

Very often                  Often                  Sometimes                  Never

b. Worked with a faculty member on activities other than coursework (committees, student groups, etc)

Very often                  Often                  Sometimes                  Never

c. Discussed course topics, ideas, or concepts with a faculty member outside of class



b. Summarized what you learned in class or from course materials  
Very often                  Often                  Sometimes                  Never

**7. During the current school year, to what extent has this course challenged you to do your best work?**

Very much                  Quite a bit                  Somewhat                  Very little                  Not at all

**8. About how many hours (per week) do you spend in a typical 7-day week preparing for class (studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, and other academic activities)?**

0                  1-5                  6-10                  11-15                  16-20                  21-25                  26-30                  More than 30

**9. Of the time you spend preparing for this course in a typical 7-day week, about how many hours (per week) are on assigned reading?**

0                  1-5                  6-10                  11-15                  16-20                  21-25                  26-30                  More than 30

**10. How much has your experience in this course contributed to your knowledge, skills, and personal development in acquiring job- or work-related knowledge and skills?**

Very much                  Quite a bit                  Somewhat                  Very little                  Not at all

**11. How would you evaluate your entire educational experience in this course?**

Excellent                  Good                  Fair                  Poor

**12. If you could start over again, would you retake this course?**

Definitely yes                  Probably yes                  Probably no                  Definitely no

**13. How likely would you be to recommend this course to another student?**

Very likely                  Somewhat likely                  Neither likely nor unlikely                  Somewhat unlikely                  Very unlikely

**14. I've learned interesting things in this course that I didn't know previously:**

Strongly agree                  Somewhat agree                  Neutral                  Somewhat disagree                  Strongly disagree

**15. I feel that the assignments and activities in this course were meaningful:**

Strongly agree                  Somewhat agree                  Neutral                  Somewhat disagree                  Strongly disagree

**16. I am proud of the work that I have done in this course:**

Strongly agree                  Somewhat agree                  Neutral                  Somewhat disagree                  Strongly disagree

**17. This course affected the way I think about my career goals:**

Strongly agree                  Somewhat agree                  Neutral                  Somewhat disagree                  Strongly disagree