



Training, Self-Efficacy, and Performance; a Replication Study

Behrooz Davazdahemami

Department of Management Science and Information
Systems, Spears School of Business
Oklahoma State University
davazda@okstate.edu

Kevin Scheibe

Department of Supply Chain & Information Systems
Iowa State University
kscheibe@iastate.edu

Andy Luse

Department of Management Science and Information
Systems, Spears School of Business
Oklahoma State University
andyluse@okstate.edu

Anthony Townsend

Department of Supply Chain & Information Systems
Iowa State University
amt@iastate.edu

Abstract:

A conceptual replication of multiple prior IS studies was conducted with the aim of providing stronger empirical support for those results. Conducting six separate longitudinal studies, the effect of professional training on improving one's application-specific computer self-efficacy (AS-CSE) was shown. Also in line with some prior IS studies it was shown that an application-specific measure of self-efficacy is better able to predict one's performance in accomplishing tasks in the corresponding domain than a general computer self-efficacy (GCSE) measure. Moreover, it is shown that, regardless of the type and characteristics of the training method, individuals' perceptions of quality of training significantly affects their AS-CSE after the training course.

Keywords: self-efficacy, training, performance, computer, application-specific.

The manuscript was received 03/07/2017 and was with the authors four months for two revisions.

1 Introduction

According to Social Cognitive Theory (Bandura, 1977), two sets of factors play major cognitive roles in guiding an individual's behavior; outcome expectations and self-efficacy. Based upon this theory, a large amount of research has been conducted in the IS literature to clarify how these factors affect or are in turn affected by one's interaction with an IT artifact (Davis, Bagozzi, & Warshaw, 1992; Hayashi, Chen, Ryan, & Wu, 2004; Henry & Stone, 1999; Hsu & Chiu, 2004; Mun & Hwang, 2003; Venkatesh, Morris, Davis, & Davis, 2003). Specifically, several studies have dealt with the relationships between self-efficacy, training, and performance in various contexts (Bouffard-Bouchard, 1990; Earley, 1994; Gist, Schwoerer, & Rosen, 1989; Stajkovic & Luthans, 1998; Tims, Bakker, & Derks, 2014).

One of the issues that has been long discussed in the IS literature is the use of either a 'general' or an 'application-specific' level of the self-efficacy construct (i.e. GSE vs. AS-SE) in the context of IT artifacts (Compeau & Higgins, 1995b; Hardin, Chang, & Fuller, 2008a, 2008b; Marakas, Johnson, & Clay, 2007, 2008). It is well discussed in prior studies that using application-specific items to capture one's self-efficacy has considerable advantages over using a general instrument for this construct (Compeau & Higgins, 1995b; Marakas et al., 2007), yet little research has specifically looked at replicating this idea across many different AS-SE measures in different IS contexts. In other words, there is still a lack of sufficient conceptual replications of this idea in the literature. In addition, research has also failed to investigate the impact of one's perception of the quality of training in forming his or her level of post-training self-efficacy.

In the present study, a conceptual replication¹ approach has been employed to replicate and re-test multiple previously addressed research questions in the IS literature to empirically validate those findings. Conducting six separate longitudinal studies on six separate groups of business college students in six different IS subject areas during several academic semesters, the present replication paper addresses the following research questions: 1) is application-specific computer self-efficacy (AS-CSE) a better predictor of performance than general computer self-efficacy (GCSE), 2) does training change trainees' level of AS-CSE over time, and 3) does trainees' perception of the quality of training affect their post-training AS-CSE? While the third issue is addressed in this study for the first time, the first two questions have been previously addressed in the IS literature by (Marakas et al., 2007) and (Torkzadeh & van Dyke, 2002) respectively. In the present study, we replicate these specific hypotheses in multiple IS contexts to provide additional empirical support for their findings. Moreover, the study entails a wide range of durations for training courses so as to investigate the effect of training duration on the findings.

We address these three questions by investigating the ability of various AS-CSE measures to assess change in AS-CSE over time as well as assessing AS-CSE in its nomological net²; by evaluating its usefulness as both a predictor of an outcome variable (performance) and as an outcome variable predicted by another exogenous variable (training quality), thereby providing a robust replication of the construct.

2 Theory and Hypotheses

Social Cognitive Theory (Bandura, 1986), is one of the most widely accepted models explaining individual behavior. According to this model there is a triadic reciprocity among environmental situations, individual cognitive and personality factors, and individual behavior; with each of these three factors either affecting or being affected by the other two. Within Social Cognitive Theory, Bandura identifies two sets of factors as the major cognitive forces guiding behavior: outcome expectations and self-efficacy (Compeau & Higgins, 1995b). He defines self-efficacy as "the people's judgments of their capabilities to organize and execute courses of actions required to attain designated types of performances" (Bandura, 1986). Self-efficacy, while associated with ability, is not a direct measure of one's real skills and abilities, but one's judgments of his or her ability to perform a task regardless of true ability.

This definition was then extended by researchers into different areas to define area-specific self-efficacy concepts like internet, academic, physical, and breastfeeding self-efficacy. (C.-L. Dennis & Faux, 1999;

¹ A conceptual replication study tests exactly the same research questions or hypotheses studied in prior research, but uses different measures, treatments, analyses, and/or context. This approach is the strongest form of replication because it ensures that there is nothing idiosyncratic about the methods and contexts employed in previous research that would limit their research conclusions (A. R. Dennis & Valacich, 2014).

² In order to help assess the validity of a developed measure, research has suggested that the measure should be evaluated by testing its relationships with other measures to which the developed measure should be related (MacKenzie et al., 2011).

Murphy, Coover, & Owen, 1989; Owen & Froman, 1988; Ryckman, Robbins, Thornton, & Cantrell, 1982; Torkzadeh & van Dyke, 2001). Also, in the IS field, based on this definition and emphasizing the distinction between the component skills someone has and his or her ability to accomplish tasks using those skills, Compeau and Higgins posited that computer self-efficacy (CSE) can be defined as the perception of ability to use computers in the accomplishment of a task (e.g. using a software package for data analysis, writing a mail merge letter using a word processor), rather than reflecting simple component skills such as formatting diskettes, booting up a computer, or using a specific software features like bolding text or changing margins. (Compeau & Higgins, 1995b).

Later, Marakas et al. (1998) proffered a more specific definition of CSE. In line with Bandura's original theory, they argued that self-efficacy is a dynamic construct that reflects more than just an ability assessment and can vary across activities and situational circumstances; therefore, a global, static, unidimensional measure of self-efficacy is not the most powerful one in predicting application-specific performance. They then advance that the specific domain of computer self-efficacy can be operationalized in three levels – general computing behavior, application specific, and application environment computer behavior – and that while the AS-CSE level refers to “the individual's perception of efficacy in performing specific computer-related tasks within the domain of general computing”, GCSE is actually the result of “the individual's judgments across multiple computer application domains.” In other words, each AS-CSE resulting from past experience carries a weight, which is unique to the individual, and all these experiences weighted together form the individual's general perception of computer self-efficacy. The authors also propose a framework for the construction of AS-CSE measuring instruments. They assert that although AS-CSE measures can be designed in many ways – with each way being more powerful than GCSE in predicting application-specific performance – they cannot be as powerful as the measure constructed with respect to the authors' proposed framework. This assertion was statistically demonstrated a few years later in a separate study (Marakas et al., 2007).

Self-efficacy, either in its general or its application-specific form, has long been discussed as one of the main antecedents of one's performance (Locke, Frederick, Lee, & Bobko, 1984; McPherson & McCormick, 2006; Multon, Brown, & Lent, 1991; Stajkovic & Luthans, 1998). However, by employing both general as well as application-specific self-efficacy instruments in a single study, Marakas et al. (2007) showed that questions eliciting estimations of ability within an application-specific rather than a general context are better predictors of performance for that specific domain. Thus, we hypothesize:

H1: An application-specific level of the self-efficacy construct is better than a more general level of this construct in predicting one's performance.

One aspect of any construct that must be understood is its stability over time, with some measures not changing while others change over time and/or across situations (MacKenzie, Podsakoff, & Podsakoff, 2011). Application-specific measures of self-efficacy have been shown to change both over time (Luse, Mennecke, & Triplett, 2013) and across situational conditions (Scheibe, Mennecke, & Luse, 2007). Using a self-developed measure of internet self-efficacy in a longitudinal study, Torkzadeh and van Dyke (2002) investigated the effect of internet training on the learners' internet self-efficacy by measuring self-efficacy before and after a training course. They showed that training significantly improves application-specific levels of self-efficacy as the trainee uses the information obtained during the course to assess abilities in performing tasks related to that domain. Although the issue was addressed in an organizational context by Earley (1994) prior to this study, we found no replication of the Torkzadeh and van Dyke study across various IS contexts nor using the AS-CSE measures developed based on the proposed framework of Marakas et al (2007). Therefore, we hypothesize:

H2: Domain-specific training significantly increases one's application-specific self-efficacy in performing tasks related to that specific domain.

Bandura, in his original Social Cognitive Theory (1986), suggests that there are four primary forces that, in order of magnitude of effect, can serve to form one's self-efficacy judgment: 1) enactive mastery (EM) which involves having the trainee actively perform the task during experiential training; 2) vicarious experiences (VE) which involves viewing another person actively performing the task; 3) verbal persuasion by someone that the individual trusts and views as competent; and 4) emotional arousal which involves one being affected by factors like stress and anxiety in her judgment about her ability.

Among these forces, it is argued that EM and VE reflect two distinct types of training conditions (Luse, Mennecke, & Townsend, 2013) as it has been employed by several researchers to design experimental

studies (Compeau & Higgins, 1995a; Gist et al., 1989). Consistent with Bandura's theory, it has been shown in these studies that EM is generally more influential than VE in forming individuals' self-efficacy. Regardless of the titles and specifications defined by Bandura for these two types of training, a general interpretation of this argument could be that different types of training differentially influence trainees' self-efficacy.

Irrespective of the training method, the subject's perception of the training quality influences self-efficacy. When one believes the domain-specific training he or she receives to be of high quality, a higher level of self-efficacy will be formed as self-efficacy is basically an assessment made by the individual which is inevitably affected by his or her judgments about all the situational factors involved (including training); in fact this is the main reason we believe it is important to study the effect of this training-related factor, among others, on self-efficacy as they both are of the same nature (i.e. self-judgment). It is also consistent with Clark et al.(1993) findings, which suggest training utility judgments (i.e. reactions to training) impact trainee motivations. Moreover, in a meta-analysis study of the training literature, Colquitt et al.(2000) hold that post-training self-efficacy is the main motivational training outcome. Hence, if individuals see higher value in the training (i.e. higher perceived quality), they will have higher judgment regarding the utility of training for helping them use technology, and this positively affects their post-training self-efficacy. We believe that it is also in line with the Bandura's theory, as a low perception of training quality serves as a low-quality vicarious experience and consequently is less able to affect SE compared to a high training quality scenario. Thus, we hypothesize:

H3: Trainees' perceptions of the quality of domain-specific training positively influences their post-training beliefs about their ability to perform domain-specific tasks (i.e. AS-CSE).

3 Formative vs. Reflective Self-Efficacy Construct

While the idea of conceptualizing CSE at the application-specific instead of the general level is widely accepted in the literature, one of the most controversial issues regarding these constructs is about their nature and the way they must be analyzed (i.e. formatively vs. reflectively). While self-efficacy and its domain-specific constructs (including CSE) were previously considered reflective constructs for several years, Marakas et al. (2007) argued that, for a number of reasons, CSE constructs (including both GCSE and AS-CSE) must be considered and validated formatively. Among the most important reasons discussed were the non-interchangeability of indicators in formative constructs (as is the case in CSE measures) as well as the necessity for the items to covary in reflective constructs (which was not the case for CSE). Later, Hardin and colleagues (2008b) in their comment paper criticized the Marakas et al. (2007) view about this issue and discussed several reasons for treating CSE constructs reflectively. They contend that since self-efficacy, according to its substantive theory, is a psychological construct, it must be considered reflectively.

It appears that the arguments by Hardin et al. have stronger theoretical grounds. Nevertheless, there was still a considerable barrier to accepting the arguments of Hardin et al. (2008b) and considering CSE as a reflective construct. For a reflective scale to be meaningful and usable for analysis, the respondents' answers to all the items are expected to covary. Before developing and using reflective constructs in this study, we must address this theoretical issue. According to Bandura's original theory (1986), self-efficacy has three main aspects: magnitude (the level of difficulty of the task), strength (strong/weak conviction regarding magnitude), and generalizability (reproducibility of expectation across situations) (Gist, 1987). To date, all of the CSE measures developed in the literature contend to be able to measure two out of these three aspects (i.e. magnitude and strength). For instance, in the case of the GCSE scale developed by Compeau and Higgins (1995b), they argue that the level of support required by the respondent to perform the task represents the magnitude of his or her GCSE and his or her conviction about that ability indicates the strength of the GCSE. For developing AS-CSE measures, although the method used to develop items is substantially different³, the same idea is used to explain the scale's ability to measure two aspects of self-efficacy.

We believe that this effort to measure two different aspects using a single measure is the most important reason to bring about the items' covariance issue noted by Marakas et al (2007). Although the idea of measuring these two aspects makes sense, it is not possible to have a single reflective measure, in accordance with classical test theory principles, to measure both aspects together; therefore, this should be done using two separate instruments where only the aspect of interest is variable and the other aspect is

³ For AS-CSE measures, each item represents a different task and the support levels are not discussed at all. In fact, this approach uses 'task difficulty' instead of 'support level' to capture the magnitude aspect of self-efficacy.

specified and fixed. For instance, if we intend to measure internet self-efficacy strength, we argue that we have to specify and fix a certain level of magnitude (i.e. to incorporate internet-related tasks at virtually the same level of difficulty⁴ as the items of the scale) and then for such a scale we can expect the item responses to covary. This is the approach employed in this study. In developing our AS-CSE scales, for each topic we used subject matter experts' opinions to make sure that all the items are at the same level of difficulty (i.e. magnitude); thus, we relied on a reflective type of analysis for our AS-CSE measures.

4 Method

To test the hypotheses, longitudinal studies were conducted with six separate groups of business college students, each group enrolled at a Midwestern university in the US. All the students enrolled in these courses were asked to participate in a pre- and a post-training survey. Table 1 indicates the number of participants who completed both surveys as well as their demographics.

For each of the domains, relying on the framework proposed by Marakas (2007), AS-CSE measures were developed by the authors. A modified semi-structured Delphi approach (Rescher, 1998) was used in developing the measures where a panel of individuals created items and then these panelists decided on

Table 1. Samples

Class	Course Title	Age Mean (SD)	Gender	Training Subject	Training Content	Measures Collected	Size	Study Period
1	Accounting Information Systems*	22.56 (3.12)	65 ♂ 69 ♀	Web Dev.	Designing a thesis submission form	AS-CSE (web) Training Quality	134	1.5-hour Class Period
2	Fundamentals of Computer Programming	21.17 (2.09)	7 ♂ 15 ♀	Program	Intro. to Programming	GCSE AS-CSE (programming) Performance (grades)	22	Semester
3	Electronic Commerce Strategy	23.65 (3.57)	61 ♂ 17 ♀	Web Dev. Server Mgmt. Database Mgmt.	Designing and Developing a Personal Music Database	AS-CSE (web) AS-CSE (server) AS-CSE (database) Training Quality	78	3-hour Class Period
4	Information Systems Infrastructure	22.78 (2.11)	55 ♂ 13 ♀	Computer Networking	Networking+	AS-CSE (server)	68	Semester
5	Project Management	24.41 (6.06)	50 ♂ 10 ♀	Project Mgmt.	PMP	GCSE AS-CSE (Project Mgmt) Performance (grades)	60	Semester
6	Management of Information Technology	22.29 (3.81)	47 ♂ 35 ♀	Virtual Worlds	SecondLife Virtual World	AS-CSE (VW)	82	2-week Course

*This was the only class offered by the Accounting Department. The rest of classes were offered by the IS department.

⁴ This difficulty level can be determined by subject matter experts (Bejar, 1981; Impara & Plake, 1998).

the best items from the corpus developed in an iterative manner until all panelists were satisfied with the remaining items. These developed scales are mentioned in the Appendix. All the participants were asked to fill out a questionnaire involving their course's corresponding AS-CSE scale prior to and after training. Training quality was also collected as part of the post-questionnaire for classes one and three. Given the short training experiences in these two courses, we decided measuring the quality of these abbreviated training experiences was needed. Also, students' final grades (on a scale of 0-100) were considered as the performance measure for testing Hypothesis 2. The final grade was deemed an appropriate performance measure given that the semester-long courses were designed around accepted certification criteria (Certified Java Associate, Project Management Professional) and that all topics covered in the course were aimed at the concepts of interest measured in the coinciding AS-CSE measure. Conversely, given that the training exercises in classes 1, 3, and 6 did not span the entire semester, the course grade would not be an accurate measure of this one portion of the course. Furthermore, given the final course grades are not used in these courses, GCSE measures in these courses are not useful. All measures collected for each class are listed in Table 1. The table also indicates the content taught in each class, sample sizes, and the time elapsed between pre- and post-training surveys.

It should be noted that we tried to choose a wide variety of classes for the study so that they represent various types of systems as well as IT experience levels. Moreover, we implemented different training lengths as we believe including both short- and long-term training (ranging from 1.5 hour to semester-long) can more robustly validate the findings of the hypotheses we are replicating.

5 Analysis and Results

5.1 Measurement Model

Partial Least Square (PLS) was used to test the measurement models. All constructs were modeled using reflective indicators. With regard to the measurement model, the composite reliability scores as well as the

Table 2. Sample Statistics

Class	Scale	Mean	SD	Alpha	AVE	CR
1	pre_AS-CSE	3.69	2.11	0.968	0.821	0.973
	po_AS-CSE	6.48	1.65	0.917	0.671	0.934
	TRNG_QTY	3.22	1.61	0.913	0.700	0.933
2	pre_AS-CSE	2.39	1.55	0.941	0.672	0.955
	po_AS-CSE	5.94	1.17	0.921	0.649	0.95
	GCSE	3.6	1.91	0.893	0.507	0.917
3	pre_AS-CSE (Server)	5.69	2.61	0.966	0.879	0.973
	po_AS-CSE (Server)	7.06	1.99	0.959	0.860	0.968
	pre_AS-CSE(Web)	5.39	2.51	0.973	0.843	0.977
	po_AS-CSE(Web)	6.79	1.97	0.979	0.874	0.982
	pre_AS-CSE(DB)	6.41	2.46	0.934	0.794	0.951
	po_AS-CSE(DB)	7.51	1.83	0.953	0.843	0.964
	TRNG_QTY	7.08	1.84	0.934	0.756	0.949
4	pre_AS-CSE	4.12	1.68	0.927	0.824	0.954
	po_AS-CSE	7.78	1.74	0.966	0.856	0.969
5	pre_AS-CSE	4	1.63	0.966	0.879	0.973
	po_AS-CSE	6	0.75	0.953	0.864	0.962
	GCSE	5.78	0.73	0.849	0.751	0.899
6	Pre_AS-CSE	3.67	1.73	0.973	0.803	0.961
	Po_AS-CSE	5.86	1.35	0.879	0.795	0.923

Chronbach's alpha for each measure were calculated and are shown in Table 2 above. The values clearly show that construct reliabilities were acceptable as they all exceed the recommended threshold of 0.7 (Nunnally, Bernstein, & Berge, 1967).

Convergent and discriminant validity of the constructs were then assessed by examining if the average variance extracted (AVE) score exceeds the recommended threshold of 0.50 for each construct (Chin, 1998; Gefen & Straub, 2005). As shown in Table 2, AVE values for all the constructs in this study are far above 0.5.

Discriminant validity of each latent construct was tested by the method recommended by Fornell and Larcker (1981). To establish discriminant validity, the squared root of AVE needs to exceed all the inter-construct correlations. Tables 3 through 8 indicate that this condition is well held across all the constructs and all the six data sets. Diagonal values in these tables represent the square root of the AVE. Overall, these results indicate that our measurement model has good discriminant and convergent validity.

Table 3. Inter-construct Correlations with Square Root of AVE along Diagonal (Class 1)			
Variables	pre_AS-CSE	po_AS-CSE	TRNG_QTY
pre_AS-CSE	0.906		
po_AS-CSE	0.294	0.819	
TRNG_QTY	-0.008	-0.417	0.837

Table 4. Inter-construct Correlations with Square Root of AVE along Diagonal (Class 2)			
Variables	pre_AS-CSE	po_AS-CSE	GCSE
pre_AS-CSE	0.820		
po_AS-CSE	0.418	0.806	
GCSE	-0.43	-0.697	0.712

Table 5. Inter-construct Correlations with Square Root of AVE along Diagonal (Class 3)							
Variable	pre_AS-CSE (Server)	po_AS-CSE (Server)	pre_AS-CSE (Web)	po_AS-CSE (Web)	pre_AS-CSE (DB)	po_AS-CSE(DB)	TRNG QTY
pre_AS-CSE (Server)	0.938						
po_AS-CSE (Server)	0.699	0.927					
pre_AS-CSE(Web)	0.737	0.602	0.918				
po_AS-CSE(Web)	0.526	0.847	0.558	0.935			
pre_AS-CSE(DB)	0.663	0.604	0.759	0.559	0.891		
po_AS-CSE(DB)	0.544	0.787	0.472	0.756	0.636	0.918	
TRNG QTY	0.119	0.466	0.221	0.557	0.188	0.491	0.869

Table 6. Inter-construct Correlations with Square Root of AVE along Diagonal (Class 4)		
Variables	pre_AS-CSE	po_AS-CSE
pre_AS-CSE	0.908	
po_AS-CSE	0.625	0.925

Variables	pre_AS-CSE	po_AS-CSE	GCSE
pre_AS-CSE	0.938		
po_AS-CSE	0.400	0.930	
GCSE	0.306	0.543	0.867

Variables	pre_AS-CSE	po_AS-CSE
pre_AS-CSE	0.896	
po_AS-CSE	0.478	0.891

5.2 Hypotheses Testing Results

To test whether AS-CSE is a better predictor of performance than GCSE (Hypothesis 1), conducting Partial Least Squares (PLS) analysis, we compared the direct standardized effect of post-training AS-CSE with that of GCSE on student performance (measured by their final grade) for groups 2 and 5. Figures 1 and 2 indicate that for both these groups, while GCSE was not a significant predictor of performance ($\beta = -0.46$ $p = 0.48$ for Class2 and $\beta = 0.04$ $p = 0.29$ for Class5) application-specific self-efficacy (post-training) was a significant antecedent ($\beta = 0.72$ $p < 0.01$ for Class2 and $\beta = 0.30$ $p < 0.01$ for Class5). Together, these results provide support for our first hypothesis. That is, in the presence of an application-specific self-efficacy measure, a general self-efficacy measure could not be considered as a good predictor for one's performance in tasks respective to that domain. This is consistent with Marakas et al. (2007) findings in this regards.

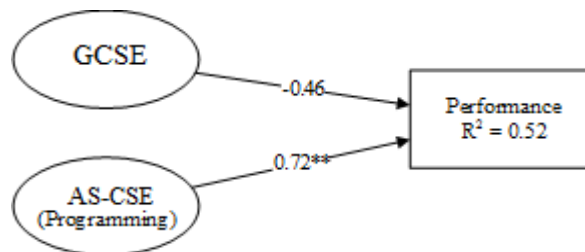


Figure 1. Impact of GCSE and AS-CSE on programming performance

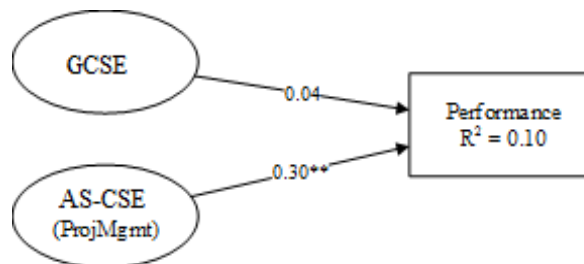


Figure 2. Impact of GCSE and AS-CSE on project management performance

To test the second hypothesis, data were analyzed to examine the pattern of application-specific self-efficacy change that participants experienced in the training courses. For each data set, mean scores across all items were calculated for both pre- and post-training AS-CSE scales. Those mean scores were then compared using a paired samples t-test to examine the significance of the difference between them. Table 9 indicates the mean scores for pre- and post-training AS-CSE scales as well as the results of the paired sample t-tests. For each pair, the mean application-specific self-efficacy score has significantly improved during the training course; providing support for Hypothesis 2.

Data Set	Pre_AS-CSE Mean (SD)	Post_AS-CSE Mean (SD)	t-value (df)	2-tail prob.	
Class 1	3.69 (2.11)	6.48 (1.45)	-17.442 (133)	<0.001	
Class 2	2.39 (1.55)	5.94 (1.17)	-12.601 (17)	<0.001	
Class 3	Server	5.69 (2.61)	7.06 (1.99)	-6.903 (77)	<0.001
	Web	5.39 (2.71)	6.79 (1.97)	-6.175 (77)	<0.001
	Database	6.41 (2.46)	7.51 (1.83)	-5.777 (77)	<0.001
Class 4	4.12 (1.68)	7.78 (1.74)	-15.51 (67)	<0.001	
Class 5	4.00 (1.63)	6.00 (0.75)	-8.901 (50)	<0.001	
Class 6	3.67 (1.73)	5.86 (1.35)	-11.118 (79)	<0.001	

To test the third hypothesis, data collected from students participating in the two shortened training exercises (Class 1 and 3) on the quality of training was used and the effect of this factor on the subjects' post-training AS-CSE was examined using PLS analysis. Figure 3 indicates the standardized direct effect of perceived quality of training on post-training AS-CSE for students enrolled in Class 1. Also, the same relationship for the subjects in Class 3 is depicted in Figure 4. As mentioned earlier, in the Class 3 dataset three AS-CSE scales were used, each for one of the skills taught in that class (i.e. Server Management, Web Developing, and Database Management). As shown in Figures 3 and 4, while the effect of quality of training on post-training AS-CSE is statistically significant in both involved data sets, this effect was negative for Class 1 but positive for Class 3 (across all its AS-CSE measures), providing partial support for Hypothesis 3 in this study. This negative relationship is discussed further in the discussion.

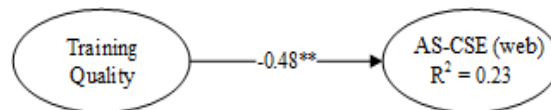


Figure 3. Impact of Training Quality on Web Development AS-CSE

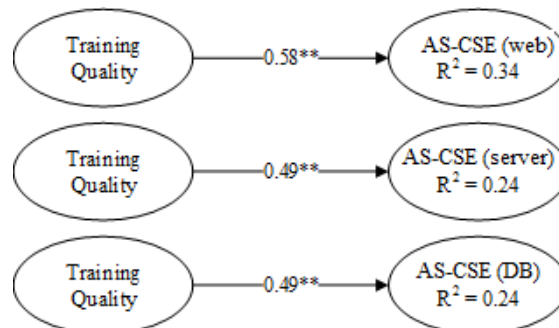


Figure 4. Impact of Training Quality on Web Development, Server, and Database AS-CSE

6 Discussion

The present study sought to address three main questions: 1) is application-specific self-efficacy a better predictor of performance than general self-efficacy, 2) does any significant change occur in individuals' level of application-specific self-efficacy after receiving domain specific training, and 3) does an individual's perception of the quality of training affect their post-training self-efficacy?

Although the first two issues were previously addressed in prior IS studies (Marakas et al., 2007; Torkzadeh & van Dyke, 2002) we conceptually replicated them with the aim of examining the validity of their findings across multiple IS contexts. Additionally, to bring clarity to the conceptual questions addressed by the mentioned two studies, we added a factor not addressed by either, specifically, the perception of the quality of the training of the task. To this end, six different longitudinal studies were conducted on six different groups of business college students taking IS courses during different time periods. The IS training provided to the trainees involved multiple computer skills including web development, programming, server management, database management, computer networking, project management, and the use of virtual worlds. Moreover, the selected courses involved a wide range of course durations ranging from single-session training to semester-long in order to increase the robustness of our findings. The research questions were investigated using data collected from the participants before and after training. Positive findings were found for each.

The first hypothesis investigated the comparative impact of application-specific versus general self-efficacy on performance. It was shown by Marakas and colleagues (2007) that an application-specific measure of the self-efficacy construct is a more powerful predictor of one's performance in accomplishing tasks than a general measure of this construct. We tested the effect of both general and application-specific self-efficacy on the performance of individuals from two separate groups in two different subject areas (programming vs. project management) at the end of training. Our results confirm those of the original study, implying that regardless of the context, participants, or situational factors, an application-specific self-efficacy measure is a more powerful predictor of performance than a general measure.

Our second hypothesis concerned the ability of training to influence application-specific self-efficacy over time. Previous research has shown (Torkzadeh & van Dyke, 2002) that the impact of training on internet skills positively affects trainees' internet self-efficacy over the duration of a course. Replicating the same experiment with six different samples of trainees trained in six different IS areas, we tested whether the results of the original study were dependent on the context, sample, or situations used by those authors. Participants in all the six studies indicated significantly higher levels of application-specific self-efficacy after the training than before. This provides additional support for the original idea of the positive effect of domain-specific training on one's application-specific self-efficacy; that is, regardless of the subject area, the people being trained, or the training settings, individuals' self-efficacy with regard to performing specific tasks can be improved by providing them with appropriate domain-specific training.

Our final hypothesis investigated the impact of training quality on the post-training measure of self-efficacy. While training type has been investigated as affecting changes in self-efficacy over time (Bandura, 1986; Compeau & Higgins, 1995a; Gist et al., 1989), the current study tested the effect of the perceived quality of training on the self-efficacy construct by treating self-efficacy as a dependent variable positioned within its nomological net. Even though the effect of training-related factors on self-efficacy has been studied well in the literature and those studies cover various IS contexts, the present paper is novel in the sense that unlike most prior research we studied this effect on the application-specific level of CSE construct instead of the GCSE. We believe that using AS-CSE instead of GCSE is more appropriate because the training participants received was at the application-specific level as well. Results show that, regardless of the training type or the context, individuals' perception of the quality of training significantly influences their application-specific self-efficacy; however, results show conflicting directions of this effect. While the effect was originally hypothesized to be positive, it turned out to be significant but negative ($\beta = -0.481$, $p < 0.001$) for one of the two study groups. It suggests that although perceived quality of training is a significant determinant of post-training AS-CSE, situational factors possibly moderate this effect. One such factor could be prior domain-specific knowledge of participants. The course subject of the two classes involved in testing this hypothesis were 'Accounting Information Systems' and 'E-commerce Strategy' (Classes 1 and 3 respectively). Although both courses were IS-related with considerable reliance on computer skills, the former was offered by the Accounting Department and hence most of the students were non-IS whereas the latter course was offered by the IS Department with the majority of students majoring in IS.

With regard to the counterintuitive findings for training with regard to class 1, Gist and Mitchell (1992) suggest that three assessment processes (namely analysis of task requirements, attributional analysis of experience, and assessment of personal and situational resources/constraints) are involved in forming one's self-efficacy through training. Relying on their theory it can be said that an individual with a higher perception of quality of training has better understood the task. They argue that for a novice trainee (which is the case with class 1 students) a better understanding of the task leads to considering higher requirements (skills and time). Then, assuming the students' personal computer skills and available resources at roughly the same level (which makes sense as they were all non-IS students in class 1), those with a better understanding of the task can be expected to consider a higher amount of effort required to produce performance comparable to the instructor's. Consequently, student judgment about their ability to perform the task would be more realistic (better calibrated). Johnson et al. (2016) claim that poorly-calibrated self-efficacy in novices leads to overconfidence. Hence, it can be said that novices with a higher perception of training quality (and consequently more realistic self-efficacy judgment) tend to be less overconfident than novices with a low perception of training quality. This explains why we found counter-intuitive results for this hypothesis with respect to class 1.

Another possible factor that might explain this counter-intuitive finding is gender. It has been shown in the literature (Busch, 1995) that women tend to show lower levels of computer self-efficacy especially when it comes to complex computer tasks. A closer look at Table 1 indicates that the proportion of women in the Accounting Information Systems class was considerably higher than that in the E-Commerce Strategy class. This higher proportion could help explain the relatively lower perceptions of self-efficacy in class1 and ultimately influence its relationship with training quality.

Regardless, quality of training was found to significantly predict self-efficacy, but future research may focus on testing the same relationship by taking into account potential moderators. While prior research has tested the effects of different types of training on self-efficacy formation (Compeau & Higgins, 1995a; Gist et al., 1989), we argued that in addition to the type of training method, personal perception of quality of training also influences self-efficacy. It implicitly suggests that practitioners, instead of manipulating training methods based purely on suggested experiential conditions (e.g. enactive mastery or vicarious experiences), may also take into account trainees' expectations of a high-quality training course to better improve their post-training self-efficacy.

In addition to the replicative value of this research, the scales developed and validated in this study for capturing individuals' application-specific self-efficacy in various contexts can be employed by other researchers for studies involving these constructs. Although Marakas et al. (2007) have proposed a process for validating formative scales developed based on their framework, these scales should be treated and analyzed reflectively provided that all the items are at the same level of difficulty. In that case, it can be expected that item responses covary, as is the case for all reflective psychological latent variables. We believe this idea can be employed in future research to address the debate in the literature over how to operationalize (i.e. formative vs. reflective) AS-CSE scales.

7 Limitations and Future Research

This research looks at two constructs – performance and training quality – within the nomological net of self-efficacy to test the construct as both a predictor and outcome variable. Several other variables are theoretically related to self-efficacy in this capacity and future research should examine these other relationships. Nevertheless, these analyses provide a good first step towards understanding self-efficacy within its nomological net.

Another limitation of this study is relatively small sample sizes in some of the classes. Our power analysis indicates that with current sample sizes, and to get a medium effect size ($r=0.3$), the power of the analyses for the six samples are respectively 0.975, 0.413, 0.866, 0.822, 0.778, and 0.841. According to this analysis, a minimum sample size of 64 is required to capture at least a medium effect size at a power of 0.8. Table 1 shows that this condition is established for all but one of the classes (i.e. class 2). However, given the significant findings for both hypothesis 1 and 2 for both classes 2 and 5, we have sufficient power to find significant results for these hypotheses and that is why we did not remove class 2 from our analyses. Yet to capture larger effect sizes, future studies need to use larger samples.

8 Conclusion

This research provides a replication study exploring application-based computer self-efficacy. Using six different measures in six different contexts across six different samples, the results show that application-specific computer self-efficacy is a better predictor of performance as compared to general computer self-efficacy, training positively affects application-specific computer self-efficacy, and training quality significantly predicts application-specific self-efficacy. These results help to further solidify the usefulness of AS-CSE measures and provide robust tests of the changing nature of self-efficacy as well as its predictive and predictable nature within a nomological net.

References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall, Inc.
- Bejar, I. I. (1981). Subject matter experts' assessment of item statistics. *ETS Research Report Series*, 1981(2).
- Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology*, 130(3), 353-363.
- Busch, T. (1995). Gender differences in self-efficacy and attitudes toward computers. *Journal of Educational Computing Research*, 12(2), 147-158.
- Chin, W. W. (1998). Issues and Opinion on Structural Equation Modeling. *MIS Quarterly*, 22(1), 1.
- Clark, C. S., Dobbins, G. H., & Ladd, R. T. (1993). Exploratory field study of training motivation influence of involvement, credibility, and transfer climate. *Group & Organization Management*, 18(3), 292-307.
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, 85(5), 678-708.
- Compeau, D. R., & Higgins, C. A. (1995a). Application of social cognitive theory to training for computer skills. *Information Systems Research*, 6(2), 118-143.
- Compeau, D. R., & Higgins, C. A. (1995b). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 189-211.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology*, 22(14), 1111-1132.
- Dennis, A. R., & Valacich, J. S. (2014). A Replication Manifesto. *AIS Transactions on Replication Research*, 1, 1-5.
- Dennis, C.-L., & Faux, S. (1999). Development and psychometric testing of the freastfeeding self-efficacy scale. *Research in Nursing and Health*, 22(5), 399-409.
- Earley, P. C. (1994). Self or group? Cultural effects of training on and performance. *Administrative Science Quarterly*, 39(1), 89-117.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Gefen, D., & Straub, D. W. (2005). A practical guide to factorial validity using PIs-Graph : Tutorial and annotated example. *Communications of the Association for Information Systems*, 16(1), 91-109.
- Gist, M. E. (1987). Self-efficacy: Implications for organizational behavior and human resource management. *Academy of Management Review*, 12(3), 472-485.
- Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy of Management Review*, 17(2), 183-211.
- Gist, M. E., Schwoerer, C., & Rosen, B. (1989). Effects of alternative training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology*, 74(6), 884-891.

- Hardin, A. M., Chang, J. C., & Fuller, M. A. (2008a). Clarifying the use of formative measurement in the IS discipline : The case of computer self-efficacy. *Journal of the Association for Information Systems*, 9, 544–546.
- Hardin, A. M., Chang, J. C., & Fuller, M. A. (2008b). Formative versus reflective measurement : Comment on Marakas , Johnson , and Clay (2007). *Journal of the Association for Information Systems*, 9(9), 519–535.
- Hayashi, A., Chen, C., Ryan, T., & Wu, J. (2004). The role of social presence and moderating role of computer self efficacy in predicting the continuance usage of e-learning systems. *Journal of Information Systems Education*, 15(2), 139-154.
- Henry, J. W., & Stone, R. W. (1999). The impacts of end-user gender, education, performance, and system use on computer self-efficacy and outcome expectancy. *Southern Business Review*, 25(1), 10-16.
- Hsu, M.-H., & Chiu, C.-M. (2004). Internet self-efficacy and electronic service acceptance. *Decision Support Systems*, 38(3), 369–381.
- Impara, J. C., & Plake, B. S. (1998). Teachers' ability to estimate item difficulty: A test of the assumptions in the Angoff standard setting method. *Journal of Educational Measurement*, 35(1), 69–81.
- Johnson, R. D., Li, Y., & Dulebohn, J. H. (2016). Unsuccessful performance and future computer self-efficacy estimations: Attributions and generalization to other software applications. *Journal of Organizational and End User Computing (JOEUC)*, 28(1), 1–14.
- Locke, E. A., Frederick, E., Lee, C., & Bobko, P. (1984). Effect of self-efficacy, goals, and task strategies on task performance. *Journal of Applied Psychology*, 69(2), 241-251.
- Luse, A., Mennecke, B. E., & Townsend, A. (2013). Experience richness: Effects of training method on individual technology acceptance. Paper presented at the 46th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- Luse, A., Mennecke, B., & Triplett, J. (2013). The changing nature of user attitudes toward virtual world technology: A longitudinal study. *Computers in Human Behavior*, 29(3), 1122–1132.
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2), 293–334.
- Marakas, G. M., Johnson, R. D., & Clay, P. F. (2007). The evolving nature of the computer self-efficacy construct: An empirical investigation of measurement construction, validity, reliability and stability over time. *Journal of the Association for Information Systems*, 8(1), 16-46.
- Marakas, G. M., Johnson, R. D., & Clay, P. F. (2008). Formative vs. reflective measurement: A Reply to Hardin , Chang , and Fuller. *Journal of the Association for Information Systems*, 9(9), 535–543.
- Marakas, G. M., Yi, M. Y., & Johnson, R. D. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward clarification of the construct and an integrative framework for research. *Information Systems Research*, 9(2), 126–163.
- McPherson, G. E., & McCormick, J. (2006). Self-efficacy and music performance. *Psychology of Music*, 34(3), 322–336.
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30-38.
- Mun, Y. Y., & Hwang, Y. (2003). Predicting the use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human-Computer Studies*, 59(4), 431–449.
- Murphy, C. A., Coover, D., & Owen, S. V. (1989). Development and validation of the computer self-efficacy scale. *Educational and Psychological Measurement*, 49(4), 893–899.
- Nunnally, J. C., Bernstein, I. H., & Berge, J. M. (1967). *Psychometric Theory*. New York: McGraw-Hill.

- Owen, S. V., & Froman, R. D. (1988). Development of a college academic self-efficacy scale. Paper Presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA, USA.
- Rescher, N. (1998). *Predicting the Future: An Introduction to the Theory of Forecasting*. SUNY Press.
- Ryckman, R. M., Robbins, M. A., Thornton, B., & Cantrell, P. (1982). Development and validation of a physical self-efficacy scale. *Journal of Personality and Social Psychology*, 42(5), 891.
- Scheibe, K. P., Mennecke, B. E., & Luse, A. (2007). The role of effective modeling in the development of self-efficacy: The case of the transparent engine. *Decision Sciences Journal of Innovative Education*, 5(1), 21–42.
- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124(2), 240-261.
- Tims, M., Bakker, A., & Derks, D. (2014). Daily job crafting and the self-efficacy–performance relationship. *Journal of Managerial Psychology*, 29(5), 490–507.
- Torkzadeh, G., & van Dyke, T. P. (2001). Development and validation of an Internet self-efficacy scale. *Behaviour & Information Technology*, 20(4), 275–280.
- Torkzadeh, G., & van Dyke, T. P. (2002). Effects of training on Internet self-efficacy and computer user attitudes. *Computers in Human Behavior*, 18(5), 479–494.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.

Appendix A: Measures

The task-specific self-efficacy measures used for the present study are as follows (unless explicitly referenced, the questions were developed by the authors using guidelines put forth by Marakas et al., 2007):

Web Development (from Scheibe et al., 2007)

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I can open a connection between a webpage and a database.
2. I can embed into a webpage a SQL command to make a database request.
3. I can embed into a script in a web page a command to read a response from a database.
4. I can close a connection to the database from a web page.
5. I understand why I need to open a connection between a web page and a database.
6. I understand why I need to embed in a web page a SQL command to make a database request.
7. I understand why I need to embed in a web page a command to read a response from a database.
8. I understand why I need close a connection to the database from a web page.

Server Management

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I can install the operating system for a web server.
2. I can setup the software needed for hosting a website.
3. I can setup an external network connection to allow users to view content hosted on a web server.
4. I can setup the necessary security settings to allow users to view content hosted on a web server.
5. I can add new content to a web server to be viewed via an internet browser.

Database Management

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I can install database server software.
2. I can create a database table.
3. I can insert data into a database.
4. I can add users to a database.
5. I can setup security access permissions for users on a database.

Computer Networking

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I can setup DNS for a network.
2. I can setup a web server for individuals to view webpages.
3. I can setup DHCP to dynamically configure IP settings for client machines.
4. I can setup an email server to send and receive email for a domain.
5. I can setup Active Directory to allow network-wide management.
6. I can setup FTP to allow for file sharing on my network.
7. I can setup a machine to provide routing of traffic between the outside world and my network.

8. I can setup a network to allow the machines to connect with each other.
9. I can install and setup a server.

Project Management

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I would be able to create tasks.
2. I would be able to assign resources.
3. I would be able to assign tasks to resources.
4. I would be able to track task completion.
5. I would be able to format and print reports of pertinent information.

Programming

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I can compile and run a java program.
2. I can utilize primitive objects in their specific situations.
3. I can iterate through code using a 'for' loop.
4. I can create a class, including methods and variables.
5. I understand the concept of inheritance.
6. I can effectively utilize UML to pictorially describe a class.
7. I can use arrays to store and use items.
8. I can construct a method, including parameters and return values.
9. I can use an 'if' statement to make decisions in my code.
10. I understand the concept of polymorphism.

Virtual Worlds *(from Luse, Mennecke, & Triplett, 2013)*

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I can change what I am looking at in Second Life.
2. I can maneuver effectively in Second Life.
3. When using Second Life to see a presentation, I can control the audio effectively.
4. When using Second Life to see a presentation, I can read the messages that are sent to me.
5. When using Second Life to see a presentation, I can read the screen showing the presentation materials.
6. When using Second Life, I can make my avatar sit.
7. When using Second Life, I can make my avatar fly.

General Computer Self-Efficacy *(from Compeau & Higgins, 1995b)*

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

1. I am able to complete the job even if there is no one around to tell me what to do as I go.

2. I am able to complete the job even if I have never used a package like it before.
3. I am able to complete the job if I have seen someone else using it before trying it myself.
4. I am able to complete the job if someone else has helped me get started.
5. I am able to complete the job if I have just the built-in help facility for assistance.

Training Quality (from Scheibe et al., 2007)

(9-point Likert scale; 1: strongly disagree, 9: strongly agree)

These questions are provided as one example of the questions used for this scale in the accounting information systems course. These same questions were modified slightly for the other two courses.

1. The professor provided me with excellent training on how to link a Web page to a database.
2. I feel like I understood the process of linking a Web page to a database after the professor trained me.
3. I probably could have learned the material better if the professor had trained me better. (*item dropped in the model specification phase*)
4. The professor provided poor training on how to link a Web page to a database. (*item dropped in the model specification phase*)
5. The professor seemed to struggle with material while providing the training. (*item dropped in the model specification phase*)
6. The professor seemed to care about whether I learned this material.
7. The training improved my understanding of how to link a Web page to a database.
8. The training from the professor helped me to understand how to effectively link a Web page to a database.

About the Authors

Behrooz Davazdahemami is currently is a PhD student in the Management Science and Information Systems Department within the Spears School of Business at Oklahoma State University. He received his B.S. and M.S. degrees both in Industrial and Systems Engineering from Isfahan University of Technology and University of Tehran, Iran. He has been published in the conference proceedings of *Hawaii International Conference on System Sciences (HICSS)*, *European Conference of Operational Research (ECOR)*, and *International Industrial Engineering Conference (IIEC)*. His current research interests involve social media, privacy, and big data analytics in healthcare. Behrooz also has some research works in operations research methods and heuristic optimization techniques in the manufacturing context.

Andy Luse received a B.A. degree in Computer Science from Simpson College, M.S. degrees in Information Assurance, Computer Engineering, Business Administration, and Psychology, and Ph.D. degrees in Human Computer Interaction, Computer Engineering, and Information Systems from Iowa State University. He is currently an Assistant Professor in Management Science and Information Systems at Oklahoma State University. Andy's research has focused on computer security, technology acceptance, and research methods. He has been published in the *Journal of Management Information Systems*, *IEEE Transactions on Visualization and Computer Graphics*, *ACM Transactions on Computing Education*, *IEEE Transactions on Education*, *Decision Sciences Journal of Innovative Education*, *Computers and Human Behavior*, and many other outlets.

Kevin Scheibe is a Kingland Business Analytics Faculty Fellow and Associate Professor of Management Information Systems at Iowa State University. His research interests include business analytics, IT privacy and security, supply chain risk, spatial decision support systems, and IT outsourcing. He is a member of the Association for Information Systems and the Decision Sciences Institute. Dr. Scheibe has published in journals such as *Decision Sciences Journal*, *European Journal of Operations Research*, *Decision Support Systems*, *Journal of Information Privacy and Security*, *Communications of the ACM*, and *Computers in Human Behavior*. He received a PhD from Virginia Polytechnic Institute and State University.

Anthony M. Townsend is an Associate Professor of MIS at Iowa State University. He received his M.S. and Ph.D. from Virginia Polytechnic Institute and State University and conducts research in collaborative systems and virtual teams. He has published in *MIS Quarterly*, *Information Systems Research*, and the *Communications of the ACM* among other venues. He is currently conducting research in collaborative systems designed to enhance a variety of organizational processes, including security.

Copyright © 2018 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints or via e-mail from ais@aisnet.org.