



# Measuring confidence engaging in computer activities at different skill levels: Development and validation of the Brief Inventory of Technology Self-Efficacy (BITS)<sup>☆</sup>

Arne Weigold<sup>a,\*</sup>, Ingrid K. Weigold<sup>b</sup>

<sup>a</sup> Department of Psychology, Kent State University at Geauga, USA

<sup>b</sup> Department of Psychology, The University of Akron, USA

## ARTICLE INFO

### Keywords:

Evaluation methodologies  
Post-secondary education  
21st century abilities  
BITS  
Computer self-efficacy

## ABSTRACT

People in educational settings are expected to effectively use increasingly complex technology. One of the most important factors in technology use is computer self-efficacy (CSE). Although frequently assessed, available measures of CSE have various issues. Additionally, there is no current measure that delineates CSE at different skill levels and specifically targets usage in applied fields. Consequently, this set of seven studies describes the development and psychometric properties of two CSE scales that address these issues: the 18-item Brief Inventory of Technology Self-Efficacy (BITS) and the six-item Brief Inventory of Technology Self-Efficacy – Short Form (BITS-SF). The first study encompassed conceptualization and item development. The second and third studies explored and confirmed the three-factor structure of the BITS (Novice, Advanced, and Expert) in samples of Mechanical Turk workers and college students, respectively. The fourth study examined the multigroup invariance of this factor structure across men and women. The fifth study showed evidence for the temporal stability of the BITS. The sixth study assessed for validity evidence for both scales. Finally, the seventh study used latent class analysis to determine the underlying classes of the BITS-SF. The uses of the BITS and BITS-SF are discussed. Both measures are available at [www.bitssurvey.com](http://www.bitssurvey.com).

## 1. Introduction

Computer-based technology is ubiquitous in, and integral to, today's societies. Consequently, the ability to use technology is essential in many settings, including education. For example, effective computer use positively impacts productivity, competency, and successful endeavors (Torkzadeh & Lee, 2003) while improving peoples' ability to engage in home-based work (Piszczek, 2017). As summarized by Peng (2017), "information technology has become one of the largest capital assets for many organizations. However, ...employees must possess the necessary computer skills to make efficient use of the computers" (p. 31). Consequently, understanding factors that contribute to individuals' effective technology use has been a continuing topic of interest across disciplines (e.g., Compeau & Higgins, 1995b; Ewwiekpaefe et al., 2018).

Of the many factors that influence technology use, one that has received much attention is computer self-efficacy (CSE). Bandura

<sup>☆</sup> We would like to thank an anonymous reviewer for the example of how skills can vary in complexity within difficulty levels.

\* Corresponding author. Kent State University at Geauga, 14111 Claridon-Troy Road, Burton, OH, 44021, United States.

E-mail address: [aweigold@kent.edu](mailto:aweigold@kent.edu) (A. Weigold).

(1986) defined self-efficacy as confidence in one's ability to successfully engage in activities to produce desired outcomes. Since self-efficacy appraisals change based on the activity (e.g., a person may have high self-efficacy for working on a team but low self-efficacy for leading it), it is necessary to delineate self-efficacy into specific areas (Bandura, 2006). One such area is CSE, or confidence working with computers and related technology (Compeau & Higgins, 1995b). Research has consistently shown CSE to relate to and/or impact various desirable outcomes, such as learning engagement and performance (e.g., Chen, 2017; Moos & Azevedo, 2009) and actual technology use (e.g., Mcilroy et al., 2007). Consequently, CSE is currently "one of the most important constructs in information systems (IS) research" (Gupta & Bostrom, 2019, p. 71).

CSE has a prominent role in educational settings. With the amount of computer-based technology in classrooms and the increased movement toward online learning (e.g., Bang & Luft, 2013; Seaman et al., 2018), both instructors and students at all education levels are expected to develop skills necessary to effectively use changing technology. CSE is a central component of educators' use of technology (Kreijns et al., 2013) and predicts the degree to which they teach computer skills to students (Siddiq & Scherer, 2016). Likewise, students' CSE predicts their intention to use technology while learning (Mcilroy et al., 2007). Consequently, educational institutions should consider CSE when making decisions regarding teacher training and classroom technology (Ferdousi, 2019; see; Siddiq & Scherer, 2016). Other researchers have noted a dearth of literature on the role of CSE in specific educational contexts, such as online learning (e.g., Alqurashi, 2016). Consequently, there remains a need for up-to-date, well-developed CSE measures to assist researchers, educators, and administrators in understanding and effectively implementing technology into the educational process.

Given the importance of CSE, many measures have been developed to assess it. However, they are of limited utility today due to outdated items, criterion contamination, and/or issues with their development (Howard, 2014; Marakas et al., 2007). Although many authors remove outdated items from available CSE measures prior to using them, these authors usually neither add new items nor revalidate the remaining items (Marakas et al., 2007), which can lead to construct underrepresentation. Additionally, none of the current measures assesses a wide range of skill levels. Educational researchers have emphasized the need for CSE measures that include separate subscales for novice and advanced skills (see Rohatgi et al., 2016; Scherer & Siddiq, 2015). Given the expectation that employees and students interact with increasingly more complex technology (see Collins & Halverson, 2010; Nelson et al., 2019), it is likely necessary to also evaluate confidence at the expert level. There is currently only one measure that assesses CSE across novice, advanced, and expert skill levels (Murphy et al., 1989), and its items are outdated (Howard, 2014). Finally, CSE measures have typically focused on either general or very specific aspects of CSE (Gupta & Bostrom, 2019), rather than targeting confidence for computer skills needed in practical usage. The latter is of special interest in educational settings, given that employees and students are often required to have both basic and complex computer skills (see Collins & Halverson, 2010; Goode, 2010; Peng, 2017).

Overall, the currently available measures of CSE have various limitations, both in general and specific to educational settings. Consequently, the current studies sought to develop two up-to-date CSE scales—one full-length measure and one screening measure—suitable for use with individuals at novice, advanced, and expert levels and focused on confidence for practical computer skills.

### 1.1. Measuring computer self-efficacy

Many CSE measures have been developed over the last 35 years (e.g., Cassidy & Eachus, 2002; Christoph et al., 1998; Compeau & Higgins, 1995b; Conrad & Munro, 2008; Gist et al., 1989; Henry & Stone, 1997; Howard, 2014; Laver et al., 2012; Marakas et al., 2007; Murphy et al., 1989; Tsai et al., 2019). As reviews of these measures exist (e.g., Gupta & Bostrom, 2019), we will not repeat their details. Instead, we focus on three prototypical measures that have been well received: the Computer Self-Efficacy Scale by Murphy et al. (1989), the untitled measure by Compeau & Higgins, 1995b, and the Computer User Self-Efficacy Scale by Cassidy and Eachus (2002). These are arguably the most popular CSE measures (see Howard, 2014; Marakas et al., 2007), and each serves as an example of one or more issues related to the use of existing CSE scales: outdated items, criterion contamination, problematic construct operationalization, and task specificity.

#### 1.1.1. Murphy et al.'s (1989) Computer Self-Efficacy Scale

The Computer Self-Efficacy Scale (Murphy et al., 1989) was one of the first CSE measures. The authors created 32 items reflecting specific actions concerning computer use. Exploratory factor analysis in a sample of students in computer courses yielded three factors: Beginning (16 items; e.g., "Adding and deleting information from a data file"), Advanced (13 items; e.g., "Using the computer to organize information"), and Mainframe (3 items; e.g., "Logging onto a mainframe computer," p. 896).

The Computer Self-Efficacy Scale (Murphy et al., 1989) is considered a strong operationalization of the CSE construct (Marakas et al., 2007). It is also one of the only CSE measures that has novice and advanced skills loading onto separate factors. However, its items are outdated (Howard, 2014). For instance, the Mainframe subscale is no longer relevant (Rainer et al., 2003), and many items on the Advanced subscale are now considered novice (Marakas et al., 2007). Outdated items are found in many CSE measures (e.g., Cassidy & Eachus, 2002; Christoph et al., 1998; Conrad & Munro, 2008; Gist et al., 1989), underscoring the need for updated scales (Marakas et al., 2007).

#### 1.1.2. Compeau & Higgins (1995b) Computer self-efficacy measure

Compeau & Higgins (1995b) developed their measure of CSE based on Bandura's (1986) assertion that self-efficacy scales should be generalizable. Consequently, Compeau and Higgins developed ten items that assessed confidence in one's ability to use unfamiliar software packages under different conditions (e.g., "If someone else had helped me get started," p. 211). Partial Least Squares models using business employees indicated a unidimensional solution that positively predicted performance, computer outcome expectations, and computer usage.

Compeau & Higgins (1995b) measure is the most used CSE scale (Marakas et al., 2007). It benefits from strong conceptualization and focus on a task (learning how to use software) that generalizes across environments. However, the authors themselves suggest that responses can be affected by individuals' learning self-efficacy and ability to imagine themselves in hypothetical scenarios (Compeau & Higgins, 1995b), which indicates criterion contamination. Additionally, specific items may be impacted by extraneous variables (Howard, 2014), such as social anxiety for those involving other people. Criterion contamination also exists in other CSE measures (e.g., Cassidy & Eachus, 2002; Conrad & Munro, 2008; Gist et al., 1989; Henry & Stone, 1997), emphasizing the need for scales that are clearly differentiated from related constructs. Finally, Compeau and Higgins' measure is specific to software use.

### 1.1.3. Cassidy and Eachus' (2002) Computer User Self-Efficacy Scale

Cassidy and Eachus (2002) developed the Computer User Self-Efficacy Scale. They considered the task-specific nature of items used in previous scales to be concerning and sought to create a more general measure. Consequently, they developed 30 items representing feelings individuals have about computers (e.g., "Using computers is something I rarely enjoy," p. 151). Analyses using college students indicated a unidimensional structure, and the measure correlated strongly and positively with computer experience.

Although the Computer User Self-Efficacy Scale was designed to account for issues in previous measures (Cassidy & Eachus, 2002), there are concerns with its construct operationalization. It is unclear to what degree this scale assesses self-efficacy, as the instructions ask respondents to rate their feelings about computers, rather than their confidence using them. At the least, many items suffer from criterion contamination (Howard, 2014). Such issues are seen in other measures such that items or full measures may not assess CSE (e.g., Christoph et al., 1998; Conrad & Munro, 2008; Henry & Stone, 1997; Howard, 2014). Additionally, Cassidy and Eachus' (2002) assertion that CSE should be assessed generally has been critiqued by other researchers (Marakas et al., 2007).

There is a continuing debate about the necessary degree of task specificity in CSE measures. For instance, unlike Cassidy and Eachus (2002), Marakas et al. (2007) emphasized the need for CSE measures to focus on task-specific activities (e.g., using spreadsheets). In practice, researchers developing CSE measures utilize different approaches depending on the measurement goal. For example, Howard (2014) developed a non-specific 12-item CSE scale modeled after general self-efficacy measures. Conversely, Tsai et al. (2019) developed their 16-item measure to specifically assess confidence in computer programming. Both approaches have been generally well received. Therefore, we conclude that the necessary degree of task specificity depends on how authors envision their measure will be used, underscoring the need for strong construct conceptualization.

## 1.2. The current studies

The available CSE measures represent a range of strengths and limitations. Some measurement aspects, such as outdated items, demonstrate the need for scales that account for current technology. Others, including criterion contamination and problematic construct operationalization, underscore the importance of item development grounded in expert knowledge. Finally, the task specificity debate emphasizes the need for measures that reflect an understanding of their target usage scenarios.

We sought to develop updated CSE measures with strong construct operationalization. Following the example of Murphy et al. (1989), we were interested in delineating novice, advanced, and expert confidence levels. Relatedly, we chose our scales' task specificity to reflect the needs of users in applied fields such that they cover a range of common practical computer skills. Consequently, we developed two scales to address these goals: one full-length measure—the Brief Inventory of Technology Self-Efficacy (BITS)—and one shorter screening measure—the Brief Inventory of Technology Self-Efficacy – Short Form (BITS-SF).

We conducted seven studies with the purpose of developing two CSE measures and assessing their psychometric properties (i.e., factor structure, temporal stability, and aspects of validity) for use in educational and other applied settings. These studies used samples from three populations and utilized best practices for developing both self-efficacy measures generally (Bandura, 2006) and CSE measures specifically (see Marakas et al., 2007). The first study describes our item development process. This study consisted of four steps: construct conceptualization (specifying the underlying construct), initial item development (creating an initial item pool), item reduction (reducing the item pool), and item refinement (developing a final list of items).

The second through fifth studies primarily focused on the BITS. The second study assessed if the BITS items resulted in a viable factor structure using exploratory factor analysis (EFA); we also developed the BITS-SF based on the final set of items determined through the EFA. The third study examined if the factor structure of the BITS from the second study could be confirmed in a different sample using confirmatory factor analysis (CFA). The fourth study assessed for multigroup invariance of the factor structure across men and women, again using CFA. Finally, the fifth study examined if the BITS subscales showed evidence of temporal stability through test-retest reliability analyses.

The sixth study investigated both the BITS and BITS-SF scores to determine if they showed evidence of convergent, discriminant, and concurrent validity. Finally, the seventh study used latent class analysis to determine if there were underlying groups delineated by the BITS-SF that corresponded to the factor structure of the BITS.

## 2. Study 1: item development

Item development consisted of four steps designed to maximize content validity: construct conceptualization, initial item development, item reduction, and item refinement. This and all subsequent studies were approved by the Institutional Review Board at the first and/or second author's institution.

### 2.1. Step 1: construct conceptualization

Strong construct conceptualization is a critical component of scale development (see MacKenzie et al., 2011). Our construct conceptualization of the BITS and BITS-SF was based on a review of the CSE literature, described above. As a result, we sought to develop two scales. The first was meant to be a short measure (< 30 items) focusing on an individual's confidence in their ability to perform computer-related actions. We were interested in actions at three levels corresponding to those developed by Murphy et al. (1989): novice, advanced, and expert. Theoretically, we based these categories on three aspects of the five-stage model of adult skill acquisition (Dreyfus, 2004). This model focuses on skill development guided by an instructor and occurring on a continuum from less to more advanced. A learner moves along the continuum by becoming familiar with the skill and gaining experience, first through generally understanding the skill and rules associated with it and later by enacting the skill intuitively in different situations (Dreyfus, 2004; Dreyfus & Dreyfus, 1980). The five stages are novice (focusing on basic features and non-contextual rules), advanced beginner (recognizing contextual aspects), competence (discriminating among, and reacting to, relevant contextual aspects), proficiency (moving beyond rules to apply different strategies), and expert (working instinctively and outside of rules; Dreyfus, 2004). This model has been used in various educational contexts (e.g., Lyon, 2015; Ogbuanya & Chukwuedo, 2017).

We were primarily interested in elucidating the first (novice), third (competence), and fifth (expert) stages (Dreyfus, 2004), although we used Murphy et al.'s (1989) terms of novice, advanced, and expert to maintain consistency with this scale. We defined "novice" actions as skills required for daily modern computer use that are less contextually situated, such as going to Internet websites. "Advanced" actions would consist of skills that necessitate some understanding of context and appropriate use but not specialized knowledge, such as using functions in spreadsheets. Finally, "expert" actions would include skills that require a substantial amount of practice or training, such as programming software applications. Our focus was on demarcating items that would fall into each level, with more advanced users also being confident in their ability to enact the lower-level skills. We expected each level to consist of items reflecting different subcategories of computer tasks, such as interacting with hardware and the Internet. We wanted these subcategories to be equally represented across the three levels. For instance, if "hardware" were a subcategory, we would include one item representing hardware in each skill level (novice, advanced, and expert). It should be noted that the actual perceived difficulty of each skill is relative to an individual's prior knowledge, and even novice-level skills vary in complexity. For example, going to Internet websites can include using Boolean logic, which we would not consider a novice skill. As a result, we paid close attention to item wording to ensure proper understanding.

We planned the second scale, the BITS-SF, to be a condensed version of the BITS (< 10 items) that could be used for swift assessment. We intended for it to be developed based on the final BITS and consist of items at all three skill levels. A total score would be calculated to provide one number reflecting a person's confidence.

During item construction, we relied on Bandura's (2006) recommendations for developing self-efficacy scales. Our scales' focus on perceived competence to engage in diverse computer tasks is consistent with Bandura's (1986) dimension of generality. We also followed Bandura's (1986) recommendation to assess magnitude (the difficulty level an individual is confident in attaining) and strength (the level of confidence one has in attaining this level), with each measure emphasizing one of these aspects. We selected a Likert scale response format for the BITS, which assesses strength and has been shown to be appropriate for self-efficacy measures (e.g., Maurer & Andrews, 2000). Since the BITS-SF was meant to be a screening tool, we focused solely on magnitude; consequently, the BITS-SF items have dichotomous response options (Yes and No; see Bandura, 1986). Finally, consistent with both Bandura's (1999) four sources of self-efficacy (mastery experiences, vicarious experiences, persuasion, and emotional states) and Dreyfus' (2004) five-stage model of skill acquisition, we developed items for which responses would most likely stem from mastery experiences, which is the most effective self-efficacy source (Bandura, 1999).

### 2.2. Step 2: initial item development

Two important aspects of item development include covering all aspects of the general domain (i.e., computer skills) and within each domain content area (i.e., novice, advanced, expert; MacKenzie et al., 2011). To this end, we based our item development on a survey of computer enthusiasts gathered via a news item posted on Tom's Hardware, a popular website for computer professionals. Interested participants followed a secure SurveyMonkey link leading to three open-ended questions regarding what computer users would feel confident doing at novice, advanced, and expert skill levels, as well as questions about their own computer use. Participants did not receive compensation.

Two hundred seventy-five participants completed the questionnaire over a two-week period. They primarily identified as expert computer users ( $n = 176$ , 64.0%), and the vast majority had helped someone else with their computer during the past month ( $n = 262$ , 95.3%).

Participants provided hundreds of useable responses to the questions about novice (588), advanced (704), and expert (582) CSE. The first author (who is an expert computer user) sorted the responses into computer skill subcategories, such as those related to hardware and file management. The first author then developed items for all non-identical actions in each subcategory, which resulted in 594 initial items.

### 2.3. Step 3: item reduction

We administered the 594 items to 86 college students at a large Midwestern university. We selected college students due to their likelihood of representing individuals primarily in the advanced skill level, but with substantial numbers in both the novice and expert

ranges. Traditional college students have been exposed to modern technology throughout most of their lives (Rashid & Asghar, 2016); however, they often do not have the specialized training expected from experts. We based our initial item reduction on four criteria: how items related to each other using principal component analysis (PCA), the average response to each item, the congruence of items with expert categories, and how representative items were of the skill subcategories (e.g., hardware). These criteria were considered concurrently, and no items were dropped solely based on component loadings. Through our item reduction process, we trimmed the 594 initial items to 53 across nine skill subcategories. The details of this process are in [Appendix S1](#) of the online supplement.

#### 2.4. Step 4: item refinement

Following the initial item reduction, we sought to refine items and potentially reduce or combine the subcategories. We examined the 53 items in a series of samples gathered from Mechanical Turk (MTurk), an online marketplace run by Amazon where people (workers) complete online tasks for monetary compensation (Iperiotis, 2010). We selected MTurk samples since they represent a different population than the college student sample used for the initial item reduction; they also likely have more advanced and expert skill levels due to seeking out Internet work. We restricted participation to those who identified as novice or novice-to-advanced computer users, who we believed would represent our conceptualization of the advanced skill level.

To arrive at the final set of items, we gathered five MTurk samples ranging in size from 82 to 142. The first sample received the full 53 items, and the resulting data were subjected to the same four criteria as the initial 594 items (except that we used EFA instead of PCA). This led to dropping some items and slightly rewriting others; the revised items were then given to a second sample. This process continued until we reached 21 items in seven subcategories (file management, hardware, Internet, networking, operating system, software, and troubleshooting), with one item each per skill level (novice, advanced, and expert). The file management items were subject to the most rewrites; we anticipated potentially removing them at a later stage. A detailed description of this process is in [Appendix S2](#) of the online supplement.

### 3. Study 2: BITS EFA and BITS-SF development

We examined the 21 items using EFA on data collected from a new sample of MTurk workers. We again restricted participation to those who were self-identified novice or novice-to-advanced computer users.

#### 3.1. Method

##### 3.1.1. Participants

Participants were 355 U.S. MTurk workers who provided useable data. Most identified as women ( $n = 224$ , 63.5%), followed by men ( $n = 128$ , 36.1%) and other ( $n = 1$ , 0.3%). Approximately two-thirds reported being in middle adulthood (30–60 years old,  $n =$

**Table 1**

Item loadings and descriptive information for the BITS exploratory factor analysis (EFA; study 2) and confirmatory factor analysis (CFA; study 3).

Item	F1: Expert EFA/ CFA	F2: Advanced EFA/ CFA	F3: Novice EFA/ CFA	<i>M</i> ( <i>SD</i> ) EFA	Range EFA	<i>M</i> ( <i>SD</i> ) CFA	Range CFA
Expert: Software	.87/.83	-.12	-.02	1.73 (1.33)	1.00–6.00	2.00 (1.32)	1.00–6.00
Expert: Networking	.84/.82	-.03	.02	1.77 (1.31)	1.00–6.00	1.92 (1.32)	1.00–6.00
Expert: Operating System	.80/.79	.01	.04	2.02 (1.36)	1.00–6.00	1.90 (1.33)	1.00–6.00
Expert: Troubleshooting	.76/.87	.06	-.04	2.10 (1.35)	1.00–6.00	2.21 (1.47)	1.00–6.00
Expert: Hardware	.75/.77	-.05	.04	1.87 (1.34)	1.00–6.00	2.03 (1.52)	1.00–6.00
Expert: Internet	.65/.55	.13	-.01	1.97 (1.34)	1.00–6.00	2.52 (1.51)	1.00–6.00
Advanced: Networking	-.15	.83/.72	.01	3.84 (1.68)	1.00–6.00	3.20 (1.78)	1.00–6.00
Advanced: Operating System	-.04	.65/.77	.10	4.47 (1.49)	1.00–6.00	3.17 (1.88)	1.00–6.00
Advanced: Hardware	.08	.63/.79	-.01	3.54 (1.73)	1.00–6.00	4.06 (1.74)	1.00–6.00
Advanced: Troubleshooting	.16	.60/.80	.03	3.36 (1.38)	1.00–6.00	3.54 (1.71)	1.00–6.00
Advanced: Software	.29	.50/.75	-.05	3.11 (1.55)	1.00–6.00	4.05 (1.72)	1.00–6.00
Advanced: Internet	.24	.42/.57	.05	3.32 (1.74)	1.00–6.00	3.80 (1.73)	1.00–6.00
Novice: Software	-.05	-.10	.80/.88	5.84 (0.52)	3.00–6.00	5.83 (0.61)	1.00–6.00
Novice: Hardware	.09	.08	.70/.83	5.75 (0.68)	1.00–6.00	5.86 (0.58)	2.00–6.00
Novice: Troubleshooting	-.03	.05	.67/.73	5.81 (0.59)	2.00–6.00	5.88 (0.43)	4.00–6.00
Novice: Internet	-.03	-.15	.64/.84	5.87 (0.53)	1.00–6.00	5.82 (0.59)	2.00–6.00
Novice: Networking	.06	.18	.43/.83	5.69 (0.77)	1.00–6.00	5.80 (0.69)	1.00–6.00
Novice: Operating System	.00	.17	.40/.68	5.69 (0.92)	1.00–6.00	5.82 (0.71)	1.00–6.00
Factor Intercorrelations							
F1	.90/.90			1.91 (1.10)	1.00–6.00	2.10 (1.15)	1.00–6.00
F2	.61/.69	.83/.87		3.61 (1.18)	1.00–6.00	3.64 (1.37)	1.00–6.00
F3	-.30/-.21	.19/.08	.74/.86	5.77 (0.45)	3.50–6.00	5.83 (0.49)	2.40–6.00

Note. Subscale scores were calculated by taking the mean of the relevant items. Factor loadings without a “/” are from the EFA only, as CFAs do not provide this information. Internal consistency estimates for the factors are listed on the diagonals of the factor intercorrelations.



233, 65.6%), followed by young adulthood (18–29 years old,  $n = 82$ , 23.1%) and older adulthood (over 60 years old,  $n = 40$ , 11.3%).

### 3.1.2. Measures

The measures consisted of demographic questions and the 21 items developed in Study 1. The 21 items were given in a pre-determined order. We began with one in the advanced category to avoid anchoring (Marakas et al., 2007) and otherwise presented the items in random order; this order was then used in all subsequent studies and the final BITS.

### 3.1.3. Procedure

Interested participants were sent to a secure Qualtrics link consisting of an introduction statement, the BITS, and the demographic questions. Upon completion, they received \$0.50.

## 3.2. Results and discussion

We conducted an EFA on the 21 items using principal axis factoring and an oblique rotation (Oblimin; Worthington & Whittaker, 2006). Factor loadings of  $\geq 0.32$  (Tabachnick & Fidell, 2007) were significant. We considered items loading at  $\geq 0.32$  on two or more factors, and/or loading on a second factor at a difference of  $\leq 0.15$ , to be cross-loading (Costello & Osborne, 2005; Worthington & Whittaker, 2006).

The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.93, and Bartlett's test was significant ( $\chi^2_{(210)} = 3335.30$ ,  $p < .001$ ). The EFA resulted in three extracted factors: Novice, Advanced, and Expert. All had eigenvalues  $> 1.00$ , which together accounted for 49.58% of the overall variance. All items loaded at  $\geq 0.32$  on their expected factor, except the novice file management item, which loaded at 0.54 on the Advanced factor. Additionally, the advanced file management item cross-loaded onto the Advanced (0.48) and Expert (0.32) factors. As we had anticipated issues with the file management items, they were removed and the EFA rerun.

The EFA on the 18 items resulted in the same three extracted factors; all had eigenvalues  $> 1.00$ , and they accounted for 50.07% of the total variance. All items loaded above 0.32 ( $\geq 0.40$ ) on only the expected factor. Each factor consisted of six items, one each for the subcategories of hardware, Internet, networking, operating system, software, and troubleshooting. Cronbach's alphas ranged from adequate to excellent, with Novice  $\alpha = 0.74$ , Advanced  $\alpha = 0.83$ , and Expert  $\alpha = .90$ . Consequently, we retained all 18 items. The final factor loadings and descriptive statistics are shown in Table 1.

### 3.3. Development of the BITS-SF

We developed the BITS-SF from the finalized 18 items. This was meant to be a short screening tool consisting of items across skill levels and subcategories to which participants would respond dichotomously (Yes or No) according to their perceived ability to engage in the skill. To this end, we chose six items from the BITS to comprise the BITS-SF: two from each skill level and one from each subcategory. We selected items that had the highest factor loadings on each of the three EFA factors; when this was not possible, we selected the second-highest factor loading for one of the items. We began the BITS-SF with an advanced item to avoid anchoring (Marakas et al., 2007) and randomly ordered the other items; we used the same order in all subsequent studies and the final BITS-SF.

## 4. Study 3: BITS confirmatory factor analysis (CFA)

We next sought to confirm the factor structure of the BITS in a sample of college students. This sample was also used in subsequent studies to assess for validity evidence for the BITS and BITS-SF scores.

### 4.1. Method

#### 4.1.1. Participants

Participants were 220 college students from a large Midwestern university who provided useable data. Approximately two-thirds identified as women ( $n = 143$ , 66.8%), followed by men ( $n = 71$ , 33.2%). Most indicated they were European American ( $n = 150$ , 70.1%), followed by African American ( $n = 35$ , 16.4%), Latin American ( $n = 10$ , 4.7%), Asian American or Biracial/Multiracial ( $n = 6$ , 2.8%, each), other ( $n = 4$ , 1.9%), and Middle Eastern/North African ( $n = 1$ , 0.5%). Ages ranged from 18 to 56 ( $M = 20.19$ ,  $SD = 4.37$ ).

#### 4.1.2. Measures

In addition to the BITS, BITS-SF, and demographics, participants completed 21 measures used in a later study to assess for validity evidence. Descriptions of all validity measures are in Table S1 of the online supplement.

#### 4.1.3. Procedure

Participants were recruited from psychology classes through the department's human subjects pool. They were directed to a secure Qualtrics link providing the informed consent, measures, and debriefing. Participants completed the demographic questions, the BITS, the counterbalanced validity measures, and, finally, the BITS-SF. They received course or extra credit.

## 4.2. Results and discussion

We assessed three solutions. The first was a single-factor model in which all items loaded on one latent factor. The second was a three-factor uncorrelated model in which the items loaded on one of three latent factors (Novice, Advanced, or Expert); we fixed the latent factor covariances to zero. The third was a three-factor correlated model, which was identical to the three-factor uncorrelated model except that the latent factors could covary. Based on the results of the EFA, we expected the three-factor correlated model to provide the best fit.

We ran all models using Mplus v. 8.0 (Muthén & Muthén, 1998 – 2017) with robust maximum likelihood estimation. We examined several goodness-of-fit indices: the robust  $\chi^2$ , the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). The robust  $\chi^2$  should be nonsignificant; however, this rarely occurs. The CFI should be  $\geq 0.900$ , with values  $\geq 0.950$  indicating good fit. The RMSEA should be  $\leq 0.100$ , with values  $\leq 0.080$  indicating adequate fit and  $\leq 0.050$  suggesting good fit (see Byrne, 2012, for a review). The SRMR should be  $\leq 0.100$  (Kline, 2016), with  $\leq 0.050$  suggesting good fit (Byrne, 2012). We compared solutions based on changes in the Akaike Information Criterion (AIC), which is appropriate for nested and non-nested models (Burnham & Anderson, 2002).

The single-factor solution was a poor fit: robust  $\chi^2_{(135)} = 1703.29, p < .001$ ; CFI = 0.000; RMSEA = 0.230 (90% CI = 0.220, 0.240); SRMR = 0.291. The three-factor uncorrelated solution was also a poor fit: robust  $\chi^2_{(135)} = 494.79, p < .001$ ; CFI = 0.766; RMSEA = 0.110 (90% CI = 0.100, 0.121); SRMR = 0.207. However, the three-factor correlated solution was generally an adequate fit: robust  $\chi^2_{(132)} = 330.45, p < .001$ ; CFI = 0.871; RMSEA = 0.083 (90% CI = 0.072, 0.094); SRMR = 0.056.

Next, we assessed models based on the AIC. Lower AICs are preferred; if the AICs of two models differ by  $> 10$ , there is no support for the one with the higher AIC (Burnham & Anderson, 2002). The three-factor correlated solution had the smallest AIC (10,203.73), which was lower than either the single-factor (11,625.17) or three-factor uncorrelated (10,349.35) AICs. Consequently, the three-factor correlated solution was the best fit. Standardized parameter estimates, which were significant at  $p < .001$ , as well as descriptive statistics, are shown in Table 1. Cronbach's alphas were strong for the Novice ( $\alpha = 0.86$ ), Advanced ( $\alpha = 0.87$ ), and Expert ( $\alpha = 0.90$ ) factors.

As expected, results from the CFA supported the Novice, Advanced, and Expert three-factor structure in college students. This is consistent with our theoretical model (Dreyfus, 2004), our construct operationalization informed by computer experts, and the EFA results based on MTurk workers.

## 5. Study 4: BITS tests of multigroup invariance

We examined the BITS for multigroup invariance across men and women using a combined sample of MTurk workers from Study 2 and college students from Study 3; this resulted in 201 (35.1%) men and 371 (64.9%) women. As men and women consistently evidence different CSE levels, which is pronounced for more advanced skills (see Hatlevik et al., 2018), we expected the two groups to show invariance for item-factor associations (configural) and item loadings (metric), but not necessarily item intercepts (scalar; see Fischer & Karl, 2019). We determined invariance by a  $\Delta\text{CFI} < 0.010$  alongside either a  $\Delta\text{RMSEA} < 0.015$  or a  $\Delta\text{SRMR} < 0.030$  (metric) or  $< 0.010$  (scalar; Chen, 2007).

The test of configural invariance showed adequate fit: robust  $\chi^2_{(264)} = 597.13, p < .001$ ; CFI = 0.910; RMSEA = 0.066 (90% CI = 0.059, 0.074); SRMR = 0.068. The goodness-of-fit indices did not substantially change for the test of metric invariance: robust  $\chi^2_{(279)} = 597.68, p < .001$ ; CFI = 0.914; RMSEA = 0.063 (90% CI = 0.056, 0.070); SRMR = 0.080. However, there were meaningful changes for the test of scalar invariance: robust  $\chi^2_{(297)} = 678.17, p < .001$ ; CFI = 0.897; RMSEA = 0.067 (90% CI = 0.060, 0.074); SRMR = 0.095. Follow-up analyses indicated scalar invariance for all Novice and Advanced items but none of the Expert items. Additionally, latent mean comparisons showed the two groups did not significantly differ for the Novice subscale (0.19,  $p = .149$ ) but that women scored lower than men on the Advanced subscale ( $-0.46, p < .001$ ). A detailed description of this study is in Appendix S3 of the online supplement.

Our results suggest the item-factor associations and factor loadings are invariant across men and women; however, the item intercepts are only invariant for the Novice and Advanced subscales. These indicate that the BITS is a viable measure with sensitivity to known demographic differences.

## 6. Study 5: BITS test-retest reliability

We assessed for evidence of test-retest reliability using subsets of the MTurk workers gathered in Study 2 who completed the BITS a second time after either two or eight weeks; 110 and 77 participants provided useable data, respectively. Intraclass correlation coefficients (ICCs) using two-way mixed effects and absolute consistency estimates were strong and significant at  $p < .001$  for all three BITS subscales after both two and eight weeks: Novice = 0.87 and 0.80, respectively; Advanced = 0.87 and 0.92; and Expert = 0.89 and 0.93. These results provide strong evidence for the temporal stability of the BITS for up to eight weeks.

## 7. Study 6: BITS and BITS-SF convergent, discriminant, and concurrent validity

Using the college student sample described in Study 3, and the measures described in Table S1 of the online supplement, we examined the BITS and BITS-SF scores for evidence of convergent, discriminant, and concurrent validity. We assessed the former two by correlating our scales with measures of general CSE; specific CSE (i.e., smartphone use, computer use for teaching, and college

student laptop use); computer anxiety, attitudes, and outcome expectations; and the broader constructs of general self-efficacy, locus of control, self-esteem, personality, and social desirability. We focused on correlation magnitude, with  $r_s = 0.20, 0.50,$  and  $0.80$  indicating small, medium, and large effect sizes, respectively (Ferguson, 2009). We examined concurrent validity by comparing how the BITS and BITS-SF related to computer skill and college major.

### 7.1. Convergent and discriminant validity

The correlations of the BITS subscales and BITS-SF total score with the validity measures are shown in Table 2. As general CSE measures often assess CSE in the advanced range, we anticipated that the general CSE measures would evidence moderate positive correlations with the BITS Advanced subscale, small-to-moderate positive correlations with the BITS-SF, and small positive correlations with the Novice and Expert subscales. Additionally, we anticipated all subscales, and the BITS-SF, would have small positive correlations with the specific CSE measures.

As expected, the BITS Advanced subscale had primarily moderate correlations with all CSE scales, with several being small ( $r$  range = 0.26 to 0.64). The BITS Novice subscale evidenced small correlations with approximately half of the CSE measures; the rest were negligible ( $r$  range =  $-0.01$  to 0.27). The BITS Expert subscale showed small correlations with all CSE measures except one, which was

**Table 2**  
Correlations of the BITS subscales and BITS-SF with validity measures in study 6.

Measure	Novice	Advanced	Expert	BITS-SF	<i>M</i> ( <i>SD</i> )	$\alpha$
<b>Convergent</b>						
Murphy et al. (1989): Comp. SE						
Beginner	.30**	.47**	.20**	.40**	4.03 (0.78)	.93
Advanced	.12	.62**	.47**	.61**	3.43 (0.88)	.92
Mainframe	.16*	.26**	.20**	.34**	3.72 (1.05)	.87
Cassidy and Eachus (2002): Comp. SE	.30**	.47**	.21**	.38**	4.32 (0.77)	.94
Compeau and Higgins (1995b): Comp. SE	.13	.48**	.36**	.44**	6.15 (1.73)	.91
Henry and Stone (1997): Comp. SE	.11	.53**	.36**	.41**	3.55 (0.78)	.79
Hill, Smith, & Mann, (1987): Comp. SE	.33**	.29**	.10	.13	3.86 (0.88)	.81
Howard (2014): Comp. SE	.13	.62**	.46**	.51**	4.48 (1.34)	.95
Marakas et al. (2007): Comp. SE						
General	.27**	.64**	.36**	.54**	6.81 (2.04)	.86
Windows	.25**	.62**	.37**	.50**	6.98 (2.23)	.84
Spreadsheet	.14*	.61**	.40**	.47**	6.15 (2.72)	.95
Word-Processing	.22**	.57**	.32**	.45**	7.01 (2.54)	.90
Internet	.26**	.57**	.30**	.51**	6.97 (2.42)	.86
Database	-.01	.37**	.38**	.39**	4.12 (2.81)	.95
Hong, Hwang, Tai, & Chen (2014): Smartphone SE	.44**	.23**	-.07	.21**	3.66 (0.48)	.84
Teo and Koh (2010): Teaching Comp. SE						
Basic	.14*	.12	-.03	.09	5.74 (1.55)	.94
Media-Related	-.03	.33**	.37**	.38**	4.26 (1.60)	.88
Web-Based	.07	.28**	.21**	.28**	5.00 (1.53)	.77
Young (2001): Laptop SE	-.12	-.06	-.00	-.01	5.01 (2.94)	.96
<b>Discriminant</b>						
Lester, Yang, & James (2005): Comp. Anxiety	-.31**	-.25**	-.05**	-.19**	2.46 (0.90)	.75
Heinssen, Glass, & Knight (1987): Comp. Anxiety	-.31**	-.39**	-.18**	-.32**	2.31 (0.63)	.90
Nickell & Pinto (1986): Comp. Attitudes	.28**	.24**	.01	.15*	3.56 (0.56)	.85
Compeau & Higgins (1995a): Outcome Expect.	.04	.29**	.23**	.20**	3.69 (0.82)	.93
Henry and Stone (1997): Outcome Expect.						
Work-Related	.18**	.30**	.12	.22**	3.88 (0.76)	.90
Personal	.15*	.23**	.12	.18*	4.07 (0.82)	.88
Chen, Gully, & Eden (2001): General SE	.30**	.14*	.07	.02	4.08 (0.55)	.90
Schwarzer & Jerusalem (1995): General SE	.14*	.21**	.10	.04	3.14 (0.44)	.87
Levenson (1973): Locus of Control						
Internal	.01	.14*	.09	.01	3.63 (0.65)	.53
Powerful Others	-.20**	.14*	.19*	.16*	2.51 (1.02)	.82
Chance	-.26**	.03	.16*	.10	2.50 (0.97)	.78
Rosenberg (1965): Self-Esteem	.18**	.10	.04	-.03	2.94 (0.53)	.88
IPIP (2020): Personality						
Extraversion	.16*	.00	-.01	-.11	3.35 (0.70)	.84
Neuroticism	-.13	-.12	-.08	-.09	2.74 (0.76)	.85
Openness	.20**	.12	-.04	-.02	3.50 (0.63)	.77
Agreeableness	.18**	-.04	-.12	-.11	3.75 (0.51)	.73
Conscientiousness	.22**	.13	.03	.07	3.44 (0.57)	.78
Hart, Ritchie, Hepper, & Gebauer (2015): Social Desirability						
Self-Deceptive Enhancement	.03	.16*	.19**	.11	4.07 (0.79)	.62
Impression Management	.10	.10	.03	.12	4.26 (0.93)	.79

Note. IPIP = International Personality Item Pool, Comp. = computer, SE = self-efficacy, Expect. = expectations. \* $p \leq .05$ , \*\* $p \leq .01$ . Bolded items are those above the minimum 0.20 effect size for practical significance.



negligible ( $r$  range = 0.10 to 0.46). The BITS-SF evidenced small-to-moderate correlations with all CSE measures except one, which was negligible ( $r$  range = 0.13 to 0.61). Finally, the three BITS subscales and BITS-SF all evidenced small correlations with half of the specialized CSE measures; the rest were negligible ( $r$  range =  $-0.00$  to 0.44).

Regarding discriminant validity, we expected the three BITS subscales and BITS-SF to evidence negligible-to-small correlations with the measures assessing computer aspects (anxiety, attitudes, and outcome expectations), as well as agency and wellbeing (general self-efficacy, locus of control, and self-esteem). We also anticipated negligible correlations with personality and social desirability. As expected, all subscales and the BITS-SF had negligible-to-small correlations with the computer anxiety, attitudes, and outcome expectation measures ( $r$  range = 0.01 to  $-0.39$ ). The BITS Advanced and Expert subscales, as well as the BITS-SF, had mostly negligible correlations with all measures of human agency, self-esteem, personality, and social desirability ( $r$  range =  $-0.01$  to 0.21). The BITS Novice subscale showed negligible correlations with all but five of these measures, which were small ( $r$  range = 0.01 to 0.30).

Overall, these results indicate good evidence of convergent and discriminant validity for the BITS subscale scores and BITS-SF total score. The BITS Novice and Expert subscales had lower correlations with other CSE measures compared to the Advanced subscale and BITS-SF. Additionally, all evidenced negligible-to-small correlations with assessments of dissimilar constructs.

## 7.2. Concurrent validity

We assessed for concurrent validity evidence by examining how the three BITS subscale scores and BITS-SF total score related to self-identified computer skill level and college major. We used multivariate analysis of variance (MANOVA; BITS) or analysis of variance (ANOVA; BITS-SF). For computer skill level, we anticipated that those who identified as novice or novice-to-advanced computer users would have significantly lower mean scores on the BITS Advanced and Expert subscales, as well as the BITS-SF, compared to those who identified as advanced or advanced-to-expert/expert users (advanced-to-expert and expert users were combined, as only three participants identified as expert users). As we expected all students to have basic computer skills, we did not hypothesize differences on the BITS Novice subscale. We made similar hypotheses regarding college major. We selected a subset of participants who were in one of two major areas that generally require strong computer skills (i.e., computer science and engineering) or in one of two areas that typically do not require advanced computer training (i.e., education and nursing). Given the results of our multigroup invariance tests in Study 4, we added gender as a control variable. We expected computer science and engineering majors to have significantly higher scores on the BITS Advanced and Expert (but not Novice) subscales, as well as the BITS-SF, than education and nursing majors.

Results are shown in Table 3. Regarding computer skill level, the resulting BITS MANOVA was significant, as were the univariate ANOVAs for the Advanced and Expert (but not Novice) subscales. The BITS-SF ANOVA was also significant. In each case, novice and novice-to-advanced users had significantly lower mean scores than advanced and advanced-to-expert/expert users. The BITS MANCOVA for major was also significant, as were the univariate ANCOVAs for the Advanced and Expert (but not Novice) subscales and the BITS-SF. Engineering majors scored significantly higher than education and nursing majors. Computer science majors only scored higher than the latter two majors for the Expert subscale; this is likely due to few participants in this major ( $n = 4$ ).

Our results show strong evidence of concurrent validity for the BITS' subscale scores and the BITS-SF total score. Subgroups expected to have higher CSE levels generally had significantly higher scores on the BITS Advanced and Expert subscales, as well as the BITS-SF, than those thought to have lower CSE levels.

**Table 3**  
BITS and BITS-SF scores for computer skill level and college major in study 6.

Comparison	<i>df</i>	<i>F</i>	<i>p</i>	$\eta^2_p$	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )
<b>Computer Skill Level</b>					<b>Novice</b>	<b>Novice/Adv</b>	<b>Adv</b>	<b>Adv/Exp</b>
BITS MANOVA	9, 630	9.08	<.001	.115				
BITS Novice	3, 210	1.59	.193	.022	5.75 (0.58)	5.86 (0.33) <sup>a</sup>	5.77 (0.68) <sup>b</sup>	6.00 (0.00) <sup>ab</sup>
BITS Advanced	3, 210	27.74	<.001	.284	2.81 (1.28) <sup>ab</sup>	3.14 (1.18) <sup>cd</sup>	4.24 (1.09) <sup>ace</sup>	5.08 (1.20) <sup>bde</sup>
BITS Expert	3, 210	15.96	<.001	.186	1.61 (1.03) <sup>ab</sup>	1.78 (0.90) <sup>cd</sup>	2.36 (1.12) <sup>ace</sup>	3.28 (1.36) <sup>bde</sup>
BITS-SF ANOVA	3, 199	12.47	<.001	.158	3.16 (0.90) <sup>ab</sup>	3.57 (1.14) <sup>cd</sup>	4.12 (0.91) <sup>ac</sup>	4.76 (1.45) <sup>bd</sup>
<b>Major</b>					<b>Education*</b>	<b>Nursing*</b>	<b>Comp Sci*</b>	<b>Engineer*</b>
BITS MANCOVA	9, 219	3.32	.001	.120				
BITS Novice	3, 73	0.60	.615	.024	5.74 (0.56)	5.82 (0.58)	6.00 (0.00)	6.00 (0.00)
BITS Advanced	3, 73	7.37	<.001	.233	2.92 (1.06) <sup>a</sup>	3.38 (1.20) <sup>b</sup>	4.42 (0.82)	5.21 (1.02) <sup>ab</sup>
BITS Expert	3, 73	7.49	<.001	.235	1.73 (0.71) <sup>ac</sup>	1.79 (0.88) <sup>bd</sup>	3.21 (1.71) <sup>ab</sup>	3.17 (1.30) <sup>cd</sup>
Gender (covariate)	3, 71	1.64	.189	.065	–	–	–	–
BITS-SF ANCOVA	3, 69	6.87	<.001	.230	3.46 (0.30) <sup>a</sup>	3.60 (0.18) <sup>b</sup>	5.23 (0.58)	5.21 (0.34) <sup>ab</sup>
Gender (covariate)	1, 69	0.01	.908	.000	–	–	–	–

Note. Novice/Adv = novice-to-advanced, Adv = advanced, Adv/Exp = advanced-to-expert/expert, Comp Sci = computer science, Engineer = engineering. BITS/BITS-SF computer skill level  $N = 214/203$ : novice  $n = 31/31$ , novice-to-advanced  $n = 94/89$ , advanced  $n = 67/62$ , advanced-to-expert/expert  $n = 22/21$ . BITS/BITS-SF college major  $N = 78/74$ : computer science  $n = 4/4$ , engineering  $n = 15/13$ , education  $n = 15/13$ , nursing  $n = 45/44$ . Means in each row with identical subscripts differ significantly at  $p \leq .05$ ; although subscripts are reported for the BITS Novice subscales, they should not be interpreted as the univariate ANOVAs were not significant. \*Standard errors are reported instead of standard deviations.

## 8. Study 7: BITS-SF classification analysis

We used latent class analysis (LCA) to assess the number of BITS-SF underlying classes. Consistent with our construct operationalization, we anticipated there would be three latent classes—novice, advanced, and expert.

We conducted LCAs using a combined sample of 391 college students from Study 3 ( $n = 204$ , 52.3%) and MTurk workers from Study 5 ( $n = 187$ , 47.8%). The college students completed the BITS-SF as part of their measure battery, and the MTurk workers did so immediately after completing the BITS for the second time. As all participants answered *Yes* to one of the novice BITS-SF items, it was removed from the analysis.

For the remaining five BITS-SF items, we estimated and compared one, two, three, and four latent classes. Of the common tests used to ascertain the optimum number, the bootstrap likelihood ratio test (BLRT) performs the strongest (Nylund et al., 2007). The BLRT uses bootstrapping to estimate differences in log likelihood tests to determine if model fit is higher for the selected number of classes compared to one fewer. A significant finding indicates the added class improves fit, whereas a nonsignificant result suggests the smaller number of classes should be selected. We also reported the AIC as a second fit test.

Both the BLRT and AIC supported the hypothesized three-class solution. This solution had significantly improved model fit compared to two classes (BLRT  $p < .001$ ), and adding a fourth class did not provide meaningful information (BLRT  $p = .429$ ). Additionally, the AIC was lowest for three classes (1581.75) compared to one (1684.44), two (1598.13), and four (1589.05) classes. These three classes fit the expected novice ( $n = 79$ , 20.2%), advanced ( $n = 248$ , 63.4%), and expert ( $n = 64$ , 16.4%) skill levels. Novice users had a high probability of responding *Yes* to the novice item ( $p = .96$ ), a low-to-moderate probability of answering *Yes* to the advanced items ( $ps = .27$  and  $.40$ ), and a low probability of answering affirmatively to the expert items ( $ps = .03$  and  $.08$ ). Advanced users had a high probability of responding affirmatively to both the novice ( $p = 1.00$ ) and advanced ( $ps = .73$  and  $1.00$ ) items, as well as a low probability of answering *Yes* to the expert items ( $ps = .01$  and  $.12$ ). Finally, expert users had a high probability of answering *Yes* to the novice ( $p = .98$ ) and advanced ( $p = .93$  and  $.95$ ) items, as well as a moderate-to-high probability of responding affirmatively to the expert items ( $ps = .62$  and  $1.00$ ).

With the final BITS-SF item added in, total scores had different ranges for novice (1–4;  $M = 2.48$ ,  $SD = 0.60$ ), advanced (3–5;  $M = 3.81$ ,  $SD = 0.52$ ), and expert (4–6;  $M = 5.45$ ,  $SD = 0.59$ ) classes. However, frequency analyses indicated that only three (3.8%) novice users and three (4.7%) expert users scored a 4; consequently, the scoring ranges are novice users 1 to 3, advanced users 3 to 5, and expert users 5 to 6. Since the categories overlap, 3 represents novice-to-advanced users, and 5 represents advanced-to-expert users; together, these correspond to the five stages of Dreyfus' (2004) model of skill acquisition.

We next examined mean scores across classes on related variables. We conducted six MANOVAs primarily involving the measures used to assess for convergent validity evidence in Study 6—one each for the three BITS subscales, one three-subscale CSE measure, five total score CSE measures, one five-subscale CSE measure, two total score specialized CSE measures, and one three-subscale specialized

**Table 4**  
Mean differences across BITS-SF classes for related variables in study 7.

Variable	Novice <i>M</i> ( <i>SD</i> )	Advanced <i>M</i> ( <i>SD</i> )	Expert <i>M</i> ( <i>SD</i> )
BITS Novice	5.69 (0.60)	5.86 (0.42)	5.71 (0.59)
BITS Advanced	2.49 <sup>a</sup> (1.00)	3.72 <sup>a</sup> (1.14)	4.56 <sup>a</sup> (1.13)
BITS Expert	1.59 <sup>a</sup> (0.81)	1.73 <sup>b</sup> (0.83)	3.48 <sup>ab</sup> (1.27)
Murphy et al. (1989): Comp. SE			
Beginner	3.54 <sup>ab</sup> (0.60)	4.12 <sup>a</sup> (0.69)	4.36 <sup>b</sup> (0.93)
Advanced	2.81 <sup>a</sup> (0.61)	3.47 <sup>a</sup> (0.79)	4.17 <sup>a</sup> (8.24)
Mainframe	3.54 <sup>ab</sup> (0.60)	4.12 <sup>a</sup> (0.69)	4.36 <sup>b</sup> (0.93)
Cassidy and Eachus (2002): Comp. SE	3.89 <sup>ab</sup> (0.60)	4.41 <sup>a</sup> (0.69)	4.67 <sup>b</sup> (0.92)
Compeau and Higgins (1995b): Comp. SE	5.27 <sup>a</sup> (1.51)	6.13 <sup>a</sup> (1.65)	7.34 <sup>a</sup> (1.69)
Henry and Stone (1997): Comp. SE	3.16 <sup>a</sup> (0.62)	3.55 <sup>a</sup> (0.72)	4.09 <sup>a</sup> (0.79)
Hill et al. (1987): Comp. SE	3.64 (0.83)	3.97 (0.81)	3.90 (1.08)
Howard (2014): Comp. SE	3.75 <sup>a</sup> (1.07)	4.57 <sup>a</sup> (1.19)	5.54 <sup>a</sup> (1.27)
Marakas et al. (2007): Comp. SE			
General	5.34 <sup>a</sup> (1.52)	6.98 <sup>a</sup> (1.86)	8.11 <sup>a</sup> (1.87)
Windows	5.13 <sup>ab</sup> (2.03)	7.25 <sup>a</sup> (1.96)	8.08 <sup>b</sup> (2.00)
Spreadsheet	4.31 <sup>ab</sup> (2.40)	6.30 <sup>b</sup> (2.57)	7.46 <sup>b</sup> (2.44)
Word-Processing	5.43 <sup>a</sup> (2.40)	7.09 <sup>a</sup> (2.48)	8.20 <sup>a</sup> (2.19)
Internet	5.28 <sup>a</sup> (1.79)	7.18 <sup>a</sup> (2.40)	8.20 <sup>a</sup> (2.19)
Database	3.00 <sup>a</sup> (1.97)	3.98 <sup>b</sup> (2.80)	5.92 <sup>ab</sup> (3.06)
Hong et al. (2014): Smartphone SE	3.43 <sup>ab</sup> (0.59)	3.72 <sup>a</sup> (0.40)	3.77 <sup>b</sup> (0.43)
Teo and Koh (2010): Teaching Comp. SE			
Basic	5.51 (1.28)	5.74 (1.72)	6.03 (1.24)
Media-Related	3.66 <sup>a</sup> (1.58)	4.20 <sup>b</sup> (1.52)	5.13 <sup>ab</sup> (1.38)
Web-Based	4.30 <sup>a</sup> (1.50)	5.04 <sup>a</sup> (1.49)	5.67 <sup>a</sup> (1.29)
Young (2001): Laptop SE	5.41 (2.42)	4.76 (2.98)	5.37 (3.39)

*Note.* Comp. = computer, SE = self-efficacy. BITS Novice, BITS Advanced, and BITS Expert means include the combined college student and MTurk worker samples. Otherwise, the sample consists only of the college students. Means with identical subscripts differ at  $p \leq .05$ . Descriptions of the measures are found in Table S1 of the online supplement.

CSE measure. The BITS subscales' MANOVA was conducted using the full sample, whereas the other MANOVAs only used the college sample (as the MTurk samples did not complete the necessary measures).

All MANOVAs and follow-up ANOVAs were significant at  $p \leq .01$ , except for one general CSE measure ( $p = .097$ ), one specific CSE measure ( $p = .308$ ), and one subscale on a specific CSE measure ( $p = .291$ ). Novice users primarily had the lowest mean scores and expert users the highest; these are shown in [Table 4](#).

Our findings support the use of the BITS-SF for assessing CSE at novice, advanced, and expert levels. Three latent classes were identified; these differed as expected on mean score comparisons across other measures.

## 9. General discussion

We conducted seven studies to develop and psychometrically assess two scales of CSE—one full-length measure (BITS) and one screening measure (BITS-SF)—that correct for issues in past scales and can be used with individuals at novice, advanced, and expert perceived computer skill levels. Both the 18-item BITS and six-item BITS-SF showed strong psychometric properties.

To emphasize content validity, we based our initial item construction on [Dreyfus' \(2004\)](#) five-stage model of adult skill acquisition and a survey of computer experts. The experts provided information on which tasks constitute novice, advanced, and expert skill levels. We also included items representing various computer-related skill subcategories (i.e., hardware, networking, Internet, operating system, software, and troubleshooting) and ensured that items from each subcategory were included at all levels to provide an appropriate degree of specificity for research and practical usage. Both the EFA and CFA supported the three-factor structure of the BITS (Novice, Advanced, and Expert), and the LCA indicated there were three classes underlying the BITS-SF that represented the same factors. The BITS' factor structure showed evidence of multigroup invariance up to the partial scalar level across men and women. Finally, both measures' scores showed good evidence of convergent, discriminant, and concurrent validity, and the BITS indicated strong temporal stability up to eight weeks.

Our scales address the limitations of existing CSE measures, which have restricted utility because of outdated items, criterion contamination, and/or problematic construct operationalization ([Howard, 2014](#)). Additionally, past measures have often been designed to be at the very general and highly specific ends of the task specificity range, rather than its center.

### 9.1. Implications

Our focus in developing the BITS and BITS-SF was to provide current measures with strong applicability to applied settings, particularly education. Most CSE measures focus on one of two types of task specificity: a general sense of CSE transcending different computer use aspects (e.g., [Howard, 2014](#)) or CSE for a specific scenario, such as using programming languages (e.g., [Tsai et al., 2019](#)). However, few existing measures allow researchers and educators to assess CSE across different computer use aspects. The BITS and BITS-SF fill these needs: They assess CSE as it represents a person's confidence for engaging in tasks at novice, advanced, and expert skill levels across different computer use subcategories that are relevant to modern technology. This is especially important for educational settings, for which CSE measures that assess confidence across different skill levels are beneficial (see [Scherer & Siddiq, 2015](#)).

The BITS and BITS-SF can be used to assess CSE in both individuals and groups and across various applications. CSE plays a prominent role in educators' use of technology ([Krejins et al., 2013](#)) and willingness to teach technology skills to their students ([Siddiq & Scherer, 2016](#)). As a result, measures such as the BITS and BITS-SF may be an important part of CSE assessment within educational settings to determine which instructors could benefit from trainings to use technology within the workplace. The BITS and BITS-SF can also be used during the hiring process to determine the highest computer skill level applicants feel confident using and if those match the complexity of technology used within the institution. Finally, the BITS and BITS-SF can also be useful for assessing student CSE. Researchers have called for educational administrators to consider individual factors, such as CSE, when determining how and when to incorporate new technology into the classroom ([Ferdousi, 2019](#)). For example, as higher education increasingly uses online student evaluations of teaching (e.g., [Chapman & Joines, 2017](#)) for their face-to-face students and continues to enroll more students in online classes ([Seaman et al., 2018](#)), it becomes essential to anticipate both how many and which students have a high likelihood of experiencing challenges using required technology. The BITS or BITS-SF can provide a quick assessment of students prior to starting college or specific classes to determine which ones may need pre-trainings or ongoing assistance.

The BITS and BITS-SF are currently the only modern measures that assess CSE across three skill levels, allowing them to be used with diverse populations for specific purposes. For example, as the BITS Novice subscale assesses confidence for basic tasks, this is useful for older adult learners, who may be uncomfortable with standard technology to the point of declining to use it ([Weigold et al., 2016](#)). Higher scores on this subscale may predict their basic classroom technology use or likelihood to enroll in online courses. The Advanced subscale would likely predict students' use of computers while learning and educators' focus on developing their students' computer skills (e.g., [Mcilroy et al., 2007](#); [Siddiq & Scherer, 2016](#)). Finally, the Expert subscale could predict students' and instructors' adaptation of new classroom technology or teaching formats, such as online learning (see [Alqurashi, 2016](#)), including the frequency and speed by which they move beyond basic functionality.

The BITS and BITS-SF are also useful for research. Educational researchers have noted a relative scarcity of literature on CSE within specific contexts and populations, such as online learning (e.g., [Alqurashi, 2016](#)). Others have emphasized the need to assess CSE across different computer skill levels to resolve contradictory past findings related to the role of teacher CSE on various outcomes (see [Scherer & Siddiq, 2015](#); [Siddiq & Scherer, 2016](#)). Our two measures' inclusion of a broad range of different computer subcategories across these levels allows them to be used with diverse research topics. Finally, their short lengths permit their inclusion in research

studies without adding undue time or cost.

The BITS-SF is also unique in that it is the first CSE measure developed as a screening tool. This is a parsimonious scale for fast assessment that can be followed up as needed with more extensive measures, including the full-length BITS. Such a scale is useful in the many situations in which educators and researchers are interested in gathering data using as few items as possible, such as when determining undergraduate students' computer confidence in Freshman orientations or including CSE items as part of the demographics of a research study.

### 9.2. Limitations

There are several directions for future research using the BITS and BITS-SF. First, both scales should be examined using samples on the low end of CSE. We selected our samples to represent a wide range of computer skill levels, centered on the advanced level. Consequently, most participants had strong perceived confidence for performing novice computer tasks. Second, our samples primarily included general workers and college students who were not chosen by random selection; further research is needed to examine these scales in various educational settings. Relatedly, examining the BITS and BITS-SF in diverse cross-cultural and adolescent samples is an important future direction, as our samples were predominantly White (college students) and in young (college students) or middle (MTurk workers) adulthood. Finally, all participants completed the studies online using computers. Although paper-and-pencil and online surveys are generally equivalent (e.g., Weigold et al., 2013), some people prefer not to complete studies on computers (see Weigold et al., 2016). Consequently, future research should examine these scales using other methods.

### 9.3. Conclusion

The present studies described the development and validation of the BITS and BITS-SF, two measures of CSE that assess individuals at three computer skill levels—novice, advanced, and expert—with items representing subcategories of computer tasks. Both scales have strong psychometric properties and address problems in prior measures; additionally, the BITS-SF is the first CSE screening measure developed. Together, these advantages make the BITS and BITS-SF ideal choices for assessing CSE. Both measures are available at [www.bitssurvey.com](http://www.bitssurvey.com) and free for non-commercial use.

### Credit author statement

Arne Weigold: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing. Ingrid Weigold: Data curation, Formal analysis, Investigation, Methodology, Resources, Visualization, Writing – original draft

### Declaration of competing interest

None.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compedu.2021.104210>.

### References

- Alqurashi, E. (2016). Self-efficacy in online learning environments: A literature review. *Contemporary Issues In Education Research*, 9(1), 45–52. <https://doi.org/10.19030/cier.v9i1.9549>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (1999). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1–45). Cambridge University Press.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares, & T. S. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Age Information Publishing.
- Bang, E., & Luft, J. A. (2013). Secondary science teachers' use of technology in the classroom during their first 5 years. *Journal of Digital Learning in Teacher Education*, 29(4), 118–126. <https://doi.org/10.1080/21532974.2013.10784715>
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus*. Routledge.
- Cassidy, S., & Eachus, P. (2002). Developing the Computer User Self-Efficacy (CUSE) Scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of Educational Computer Research*, 26(2), 133–153. <https://doi.org/10.2190/JGJR-0KVL-HRF7-GCNU>
- Compeau, D. R., & Higgins, C. A. (1995a). Application of social cognitive theory to training for computer skills. *Information Systems Research*, 6(2), 118–143. <https://doi.org/10.1287/isre.6.2.118>
- Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, 4(1), 62–83. <https://doi.org/10.1177/109442810141004>
- Chapman, D. D., & Joines, J. A. (2017). Strategies for increasing response rates for online end-of-course evaluations. *International Journal of Teaching and Learning in Higher Education*, 29(1), 47–60.

- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Chen, I.-S. (2017). Computer self-efficacy, learning performance, and the mediating role of learning engagement. *Computers in Human Behavior*, 72, 362–370. <https://doi.org/10.1016/j.chb.2017.02.059>
- Christoph, R. T., Schoenfeld, G. A., Jr., & Tansky, J. W. (1998). Overcoming barriers to training utilizing technology: The influence of self-efficacy factors on multimedia-based training receptiveness. *Human Resource Development Quarterly*, 9(1), 25–38. <https://doi.org/10.1002/hrdq.3920090104>
- Collins, A., & Halverson, R. (2010). The second educational revolution: Rethinking education in the age of technology. *Journal of Computer Assisted Learning*, 26(1), 18–27. <https://doi.org/10.1111/j.1365-2729.2009.00339.x>
- Compeau, D. R., & Higgins, C. A. (1995b). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19(2), 189–211. <https://doi.org/10.2307/249688>
- Conrad, A. M., & Munro, D. (2008). Relationships between computer self-efficacy, technology, attitudes and anxiety: Development of the computer technology use sale (CTUS). *Journal of Educational Computing Research*, 39(1), 51–73. <https://doi.org/10.2190/EC.39.1.d>
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, 10. <https://doi.org/10.7275/jyj1-4868>. Article 7.
- Dreyfus, S. E. (2004). The five-stage model of adult skill acquisition. *Bulletin of Science, Technology & Society*, 24(3), 177–181. <https://doi.org/10.1177/0270467604264992>
- Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. California University Berkeley Operations Research Center. <https://doi.org/10.21236/ada084551>
- Evwiekpaefe, A. E., Chiemeke, S. C., & Haruna, M. Z. (2018). Individual and organizational acceptance of technology theories and models: Conceptual gap and possible solutions. *Pacific Journal of Science and Technology*, 19(1), 189–197. [http://www.akamaiuniversity.us/PJST19\\_1\\_189.pdf](http://www.akamaiuniversity.us/PJST19_1_189.pdf)
- Ferdousi, B. (2019). The effect of computer self-efficacy and attitude on undergraduate students' intention to use emerging technology in classroom learning. *Journal of Computer Sciences and Applications*, 7(1), 50–55. <https://doi.org/10.12691/jcsa-7-1-8>
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532–538. <https://doi.org/10.1037/a0015808>
- Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibility in R. *Frontiers in Psychology*, 10, e1507. <https://doi.org/10.3389/fpsyg.2019.01507>
- 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. <https://doi.org/10.1037/0021-9010.74.6.884>
- Goode, J. (2010). Mind the gap: The digital dimension of college access. *The Journal of Higher Education*, 81(5), 583–618. <https://doi.org/10.1080/00221546.2010.11779068>
- Gupta, S., & Bostrom, R. P. (2019). A revision of computer self-efficacy conceptualizations in information systems. *ACM SIGMIS - Data Base: The Database for Advances in Information Systems*, 50(2), 71–93. <https://doi.org/10.1145/3330472.3330478>
- Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *Sage Open*, 5(4), 1–9. <https://doi.org/10.1177/2158244015621113>
- Hatlevik, O. E., Thronsen, I., Loi, M., & Gudmundsdottir, G. B. (2018). Students' ICT self-efficacy and computer and information literacy: Determinants and relationships. *Computers & Education*, 118, 107–119. <https://doi.org/10.1016/j.compedu.2017.11.011>
- Henry, J. W., & Stone, R. W. (1997). The development and validation of computer self-efficacy and outcome expectancy scales in a nonvolitional context. *Behavior Research Methods, Instruments, & Computers*, 29(4), 519–527. <https://doi.org/10.3758/BF03210603>
- Heinssen, R. K., Glass, C. R., & Knight, L. A. (1987). Assessing computer anxiety: Development and validation of the Computer Anxiety Rating Scale. *Computers in Human Behavior*, 3(1), 49–59. [https://doi.org/10.1016/0747-5632\(87\)90010-0](https://doi.org/10.1016/0747-5632(87)90010-0)
- Howard, M. C. (2014). Creation of a computer self-efficacy measure: Analysis of internal consistency, psychometric properties, and validity. *Cyberpsychology, Behavior, and Social Networking*, 17(10), 677–681. <https://doi.org/10.1089/cyber.2014.0255>
- Hill, T., Smith, N. D., & Mann, M. F. (1987). Role of efficacy expectations in predicting the decision to use advanced technologies: The case of computers. *Journal of Applied Psychology*, 72(2), 307–313. <https://doi.org/10.1037/0021-9010.72.2.307>
- Hong, J.-C., Hwang, M.-Y., Tai, K.-H., & Chen, Y.-L. (2014). Using calibration to enhance students' self-confidence in English vocabulary learning relevant to their judgment of over-confidence and predicted by smartphone self-efficacy and English learning anxiety. *Computers & Education*, 72, 313–322. <https://doi.org/10.1016/j.compedu.2013.11.011>
- Iperiotis, P. (2010). *Demographics of mechanical Turk*. <https://archivefda.dlib.nyu.edu/jspui/bitstream/2451/29585/2/CeDER-10-01.pdf>
- International Personality Item Pool. (2020). *A scientific collaboratory for the development of advanced measures of personality and other individual differences*. [https://ipip.ori.org/New\\_IPIP-50-item-scale.htm](https://ipip.ori.org/New_IPIP-50-item-scale.htm)
- Kreijns, K., Vermeulen, M., Kirschner, P. A., van Buuren, H., & Van Acker, F. (2013). Adopting the integrative model of behaviour prediction to explain teachers' willingness to use ICT: A perspective for research on teachers' ICT usage in pedagogical practices. *Technology, Pedagogy and Education*, 21(1), 55–71. <https://doi.org/10.1080/1475939X.2012.754371>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th). Guilford Press.
- Laver, K., George, S., Ratcliffe, J., & Crotty, M. (2012). Measuring technology self-efficacy: Reliability and construct validity of a modified computer self efficacy scale in a clinical rehabilitation setting. *Disability & Rehabilitation*, 34(3), 220–227. <https://doi.org/10.3109/09638288.2011.593682>
- Lyon, L. J. (2015). Development of teaching expertise viewed through the Dreyfus model of skill acquisition. *The Journal of Scholarship of Teaching and Learning*, 15(1), 88–105. <https://doi.org/10.14434/josotl.v15i1.12866>
- Lester, D., Yang, B., & James, S. (2005). A short computer anxiety scale. *Perceptual and Motor Skills*, 100(3 Pt 2), 964–968. <https://doi.org/10.2466/pms.100.3c.964-968>
- Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology*, 41(3), 397–404. <https://doi.org/10.1037/h0035357>
- 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. <https://doi.org/10.2307/23044045>
- 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. <https://doi.org/10.17705/1jais.00112>
- Maurer, T. J., & Andrews, K. D. (2000). Traditional, Likert, and simplified measures of self-efficacy. *Educational and Psychological Measurement*, 60(6), 965–973. <https://doi.org/10.1177/00131640021970899>
- McIlroy, D., Sadler, C., & Boojawon, N. (2007). Computer phobia and computer self-efficacy: Their association with undergraduates' use of university computer facilities. *Computers in Human Behavior*, 23, 1285–1299. <https://doi.org/10.1016/j.chb.2004.12.004>
- Moos, D. C., & Azevedo, R. (2009). Learning with computer-based learning environments: A literature review of computer self-efficacy. *Review of Educational Research*, 79(2), 576–600. <https://doi.org/10.3102/0034654308326083>
- 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. <https://doi.org/10.1177/001316448904900412>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Nelson, M. J., Voithofer, R., & Cheng, S.-L. (2019). Mediating factors that influence the technology integration practices of teacher educators. *Computers & Education*, 128, 330–344. <https://doi.org/10.1016/j.compedu.2018.09.023>
- Nickell, G. S., & Pinto, J. N. (1986). The computer attitude scale. *Computers in Human Behavior*, 2(4), 301–306. [https://doi.org/10.1016/0747-5632\(86\)90010-5](https://doi.org/10.1016/0747-5632(86)90010-5)



- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes of latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535–569. <https://doi.org/10.1080/10705510701575396>
- Ogbuanya, T. C., & Chukwuodo, S. O. (2017). Career-training mentorship intervention via the Dreyfus model: Implication for career behaviors and practical skills acquisition in vocational electronic technology. *Journal of Vocational Behavior*, 103, 88–105. <https://doi.org/10.1016/j.jvb.2017.09.002>
- Peng, G. (2017). Do computer skills affect worker employment? An empirical study from CPS surveys. *Computers in Human Behavior*, 74, 26–34. <https://doi.org/10.1016/j.chb.2017.04.013>
- Piszczek, M. M. (2017). Boundary control and controlled boundaries: Organizational expectations for technology use at the work-family interface. *Journal of Organizational Behavior*, 38(4), 592–611. <https://doi.org/10.1002/job.2153>
- Rainer, R. K., Jr., Laosethakul, K., & Astone, M. K. (2003). Are gender perceptions of computing changing over time? *Journal of Computer Information Systems*, 43(4), 108–114.
- Rashid, T., & Asghar, H. M. (2016). Technology use, self-directed learning, student engagement and academic performance: Examining the interrelations. *Computers in Human Behavior*, 63, 604–612. <https://doi.org/10.1016/j.chb.2016.05.084>
- Rohatgi, A., Scherer, R., & Hatlevik, O. E. (2016). The role of ICT self-efficacy for students' ICT use and their achievement in a computer and information literacy test. *Computers & Education*, 102, 103–116. <https://doi.org/10.1016/j.compedu.2016.08.001>
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.
- Scherer, R., & Siddiq, F. (2015). Revisiting teachers' computer self-efficacy: A differentiated view on gender differences. *Computers in Human Behavior*, 53, 48–57. <https://doi.org/10.1016/j.chb.2015.06.038>
- Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). NFER-NELSON.
- Seaman, J. E., Allen, I. E., & Seaman, J. (2018). *Grade increase: Tracking distance education in the United States*. Babson Survey Research Group. <http://www.onlinelearningsurvey.com/highered.html>.
- Siddiq, F., & Scherer, R. (2016). The relation between teachers' emphasis on the development of students' digital information and communication skills and computer self-efficacy: The moderating roles of age and gender. *Large-scale Assessments in Education*, 4(1), 17. <https://doi.org/10.1186/s40536-016-0032-4>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5<sup>th</sup> ed.). Allyn & Bacon/Pearson Education.
- Teo, T., & Koh, J. H. L. (2010). Assessing the dimensionality of computer self-efficacy among pre-service teachers in Singapore: a structural equation modeling approach. *International Journal of Education and Development using Information and Computer Technology*, 6(3), 7–18.
- Torkzadeh, G., & Lee, J. (2003). Measures of perceived end-user computing skills. *Information & Management*, 40(7), 607–615. [https://doi.org/10.1016/S0378-7206\(02\)00090-3](https://doi.org/10.1016/S0378-7206(02)00090-3)
- Tsai, M.-J., Wang, C.-Y., & Hsu, P.-F. (2019). Developing the computer programming self-efficacy scale for computer literacy education. *Journal of Educational Computing*, 56(8), 1345–1360. <https://doi.org/10.1177/0735633117746747>
- Weigold, A., Weigold, I. K., Drakeford, N. M., Dykema, S. A., & Smith, C. A. (2016). Equivalence of paper-and-pencil and computerized self-report surveys in older adults. *Computers in Human Behavior*, 54, 407–413. <https://doi.org/10.1016/j.chb.2015.08.033>
- Weigold, A., Weigold, I. K., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based paper-and-pencil and Internet data collection methods. *Psychological Methods*, 18(1), 53–70. <https://doi.org/10.1037/a0031607>
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806–838. <https://doi.org/10.1177/0011000006288127>
- Young, M. R. (2001). Windowed, wired, and webbed—now what? *Journal of Marketing Education*, 23(1), 45–54. <https://doi.org/10.1177/0273475301231006>