J Sci Educ Technol (2015) 24:696–708
DOI 10.1007/s10956-015-9558-4

Scientific Inquiry Self-Efficacy and Computer Game Self-Efficacy as Predictors and Outcomes of Middle School Boys' and Girls' Performance in a Science Assessment in a Virtual Environment

Bradley W. Bergey[1] · Diane Jass Ketelhut[2] · Senfeng Liang[3] · Uma Natarajan[4] · Melissa Karakus[5]

Published online: 26 March 2015

Abstract The primary aim of the study was to examine whether performance on a science assessment in an immersive virtual environment was associated with changes in scientific inquiry self-efficacy. A secondary aim of the study was to examine whether performance on the science assessment was equitable for students with different levels of computer game self-efficacy, including whether gender differences were observed. We examined 407 middle school students' scientific inquiry self-efficacy and computer game self-efficacy before and after completing a computer game-like assessment about a science mystery. Results from path analyses indicated that prior scientific inquiry self-efficacy predicted achievement on end-of-module questions, which in turn predicted change in scientific inquiry self-efficacy. By contrast, computer game self-efficacy was neither predictive of nor predicted by performance on the science assessment. While boys had higher computer game self-efficacy compared to girls, multi-group analyses suggested only minor gender differences in how efficacy beliefs related to performance. Implications for assessments with virtual environments and future design and research are discussed.

Keywords Self-efficacy · Scientific inquiry · Computer games · Immersive virtual environments · Gender differences

✉ Bradley W. Bergey
bradley.bergey@dal.ca

[1] Department of Psychological and Neuroscience, Dalhousie University, 1355 Oxford Street, P. O. Box 15000, Halifax, NS B3H 4R2, Canada

[2] Teaching and Learning, Policy and Leadership, University of Maryland, College Park, MD, USA

[3] University of Wisconsin—Stevens Point, Stevens Point, WI, USA

[4] Education Development Center, Waltham, MA, USA

[5] Department of Psychological, Organizational and Leadership Studies, Temple University, Philadelphia, PA, USA

Introduction

The manner in which science knowledge and inquiry skills are evaluated has implications for how students perceive their scientific abilities. These perceptions predict future achievement in science (Multon et al. 1991) and shape career aspirations and trajectories (Bandura et al. 2001). For middle school science students, performance on science tests is an important source of information about whether he or she is capable of excelling in the domain. Since this is a key age for science career choices (Tai et al. 2006), it is crucial that middle school students receive reliable and valid information about these capabilities. Unfortunately, too often traditional assessments focus on a limited set of content-based skills, ignoring scientific inquiry skills (National Research Council 2005). Immersive virtual environments (IVEs)—computer-based worlds in which individuals direct characters to solve contextually situated challenges—are a promising approach to science assessment for several reasons; chief among these is that IVEs can assess the application of scientific knowledge and inquiry skills to solve contextualized problems in ways that resemble how scientists actually work (Clark et al. 2009; Clarke-Midura and Dede 2010; Ketelhut et al. 2013; Timms et al. 2012). Important questions arise from situating assessment activities that require enacting scientific inquiry skills within a virtual environment. One question is whether the format of the assessment is fair for students with different computer playing abilities, given that some science assessments in IVEs have computer game-like features. A second question is whether experiencing success or failure on a science assessment is associated with students' beliefs, not just their skills, about their ability to conduct scientific inquiry. To address these questions, the current study examines whether a science assessment in a virtual environment is fair for boys and girls who may differ in their perceived ability to succeed at computer games and whether performance on a science assessment in an IVE is associated with changes in scientific inquiry ability beliefs.

Self-Efficacy

Self-efficacy refers to the self-belief that one is able to perform specific tasks or processes and to achieve designated results (Pajares 1996). In his seminal work on self-efficacy, Bandura (1977) argued that a person's belief that he or she is capable of succeeding at a task strongly influenced effort expenditure and persistence on the task. Thus, the belief that one can be successful helps individuals regulate their behaviors, with higher self-efficacy associated with higher levels of effort, engagement, and persistence (Pajares 1996). Much empirical research has demonstrated the positive and often powerful effect of self-efficacy on achievement (for reviews see Pajares 1996; Stajovic and Luthans 1998; Wigfield and Eccles 2002; Zimmerman 2000).

Efficacy beliefs are not static (Bandura 1997); they change in response to experiences and environmental and psychological factors. Bandura (1986, 1997) theorized four sources of individuals' self-efficacy beliefs: mastery experiences (e.g., prior success or failure with a similar task), vicarious experiences (e.g., watching peers), social persuasion (e.g., encouragement), and physiological feedback (e.g., stress). Of these sources, mastery experiences—an individual's perceived prior mastery with similar tasks—are the most powerful influence (Bandura 1986, 1997). For example, Britner and Pajares (2006) found that mastery experiences were the strongest source in predicting science self-efficacy beliefs among middle school students. Therefore, as students perceive experiences of success or failure with tasks, they may change their self-efficacy beliefs for related tasks accordingly (for moderating influences, see Schunk 1983; Weiner 2005).

Self-efficacy beliefs are not global but rather are specific to the demands of a given task (Bandura 1986). In the context of science assessments based in a computer game-like immersive virtual environment, the task demands include both doing scientific inquiry and manipulating the virtual environment in which the science inquiry task is placed. Accordingly, in the current study, we examine *scientific inquiry self-efficacy,* which we define as the self-belief that one is capable of performing the tasks involved in scientific inquiry, and *computer game self-efficacy,* which we define as the self-belief that one is capable of performing computer game tasks (Ketelhut 2010). Investigating the two simultaneously allows us to explore the impact the format of this assessment has on students' ability to be successful. A potential finding that computer game self-efficacy predicts student success would call into question the validity of these types of assessment.

Scientific Inquiry Self-Efficacy

In the context of middle school science, researchers have found that science self-efficacy positively predicts academic achievement (Britner and Pajares 2001; Wang et al. 2007). A small number of studies have found that a type of science self-efficacy—*scientific inquiry* self-efficacy—is predictive of performance on computer game-like IVEs that require scientific inquiry (Chen et al. 2014; Clark et al. 2009; Ketelhut 2007; Nelson and Ketelhut 2008). For example, in the IVE River City in which players must solve the problem of why people of a late 1800s town are getting sick, scientific inquiry self-efficacy beliefs prior to playing the module predicted early inquiry behaviors (Ketelhut 2007) and adaptive use of an in-module guidance system (Nelson and Ketelhut 2008).

Results from a small number of studies examining whether scientific inquiry self-efficacy beliefs change in response to experiences with science inquiry tasks in IVEs are mixed and raise questions for future research. Ketelhut (2007) investigated inquiry behaviors over time in an IVE (River City) and found no difference in students' self-efficacy beliefs before and after the experience. By contrast, Chen et al. (2014) observed increases in scientific inquiry self-efficacy among fifth graders after participating in a 10-day curriculum that involved an IVE about causal relations in a marine ecosystem. Similarly, Meluso et al. (2012) examined the science self-efficacy (as opposed to scientific inquiry self-efficacy) of fifth grade students before and after they participated in an IVE computer game about landforms and found that students' science self-efficacy increased after playing. None of the above studies examined change in self-efficacy beliefs in relation to level of achievement in the IVE, only in relation to participation. Presumably, the majority of students in these studies had mastery experiences in the IVEs, thereby leading to group-level increases in self-efficacy beliefs. When IVEs are designed to assess knowledge with a challenging assessment task over a relatively brief period of time rather than assist learning over repeated engagement, the extent to which students attain mastery may be more variable. In such situations, changes in self-efficacy beliefs may be observed at the level of the individual but not of the group, though to our knowledge this prediction has not been tested.

Computer Game Self-Efficacy

Computer game self-efficacy may influence performance in IVEs when the learning or assessment task in the IVE shares features with computer games. Educational IVEs are often developed using principles of game design (Nelson and Erlandson 2012). Like many computer games, players in educational IVEs often need to understand a goal or mission, direct an avatar (virtual personal character) through visually and aurally complex worlds, scout virtual terrain, interact with objects and characters, experiment with possible actions (e.g., jumping, using tools, etc.), and synthesize these actions in order to solve narrative-based problems. Given these computer game-like characteristics of some IVEs, prior experiences and self-beliefs about one's ability to succeed at computer games may relate to performance in educational IVEs. Specifically, students with greater computer game self-efficacy may exert more effort and persistence within the IVE, resulting in higher scores on the assessment. Thus, it is reasonable to hypothesize that computer game self-efficacy may influence achievement in computer game-like IVEs.

Empirical research on the relations between computer game self-efficacy and performance in educational IVEs is scarce (Pavlas et al. 2010). Recent comprehensive review articles (Clark et al. 2009; Mikropoulos and Natsis 2011) and a meta-analysis (Wouters et al. 2013) in the field of educational gaming made no mention of research examining computer game self-efficacy. In our review of the literature on educational games, we were unable to locate any research articles that examined computer game self-efficacy or related concepts at the middle school level. A few studies have been conducted with *video game* self-efficacy with undergraduate students, which report findings that indicate that video game self-efficacy and prior experience with video games are important predictors of performance in video games (e.g., Orvis et al. 2008; Pavlas et al. 2010). For example, Pavlas et al. reported that video game self-efficacy predicted undergraduate students' learning and intrinsic motivation while playing an educational IVE about the immune system. The broader literature on video game playing in general has identified efficacy as an important predictor of choice and performance in video games (e.g., Klimt and Hartmann 2006). In addition, related forms of self-efficacy—such as technology self-efficacy—have predicted success in e-learning environments (e.g., Johnson et al. 2008). Thus, one potential problem with using virtual environments as an assessment is that such environments could privilege students who are efficacious in playing computer games. One of the goals of the current study was to examine the relations between computer game self-efficacy and performance in an educational assessment IVE.

Gender Differences

The social and psychological influences that produce gender inequality in science, technology, engineering, and mathematics (STEM) fields are many and complex (Hill et al. 2010). Gender differences in self-efficacy beliefs have been identified as one important variable (Lapan et al. 1996). Gender differences in science self-efficacy can result from many influences. Educational, cultural, mass media, and home influences contribute to adolescents' conceptualization of certain academic domains or careers (e.g., STEM) as being gender stereotyped; as a result, boys and girls may develop values and ability beliefs for achievement in those domains accordingly (for a broader discussion, see Bussey and Bandura 1999; Eccles et al. 1983; Pajares 2002). Among adolescents, however, empirical studies on gender differences in self-efficacy beliefs in gendered domains like STEM have been mixed (Schunk and Meece 2005) although recent research examining gender differences in science self-efficacy among middle school students has found no significant differences (Britner and Pajares 2001, 2006; Chen and Usher 2013; Kiran and Sungur 2012; Usher and Pajares 2008). The current study examines potential gender differences for a particular type of science self-efficacy belief: scientific inquiry self-efficacy. To our knowledge, no prior research has investigated this comparison. Thus, one goal of the current study is to examine whether middle school boys and girls exhibit different scientific inquiry self-efficacy beliefs and whether these beliefs have different relationships with performance in the IVE.

Despite ongoing changes in the type and availability of video and computer games, researchers consistently find that males report liking and playing video games more than females (e.g., Buchman and Funk 1996; Hartmann and Klimmt 2006; Lucas and Sherry 2004; Nietfeld et al. 2014; Terlecki et al. 2011; Turkle 1995; Wright et al. 2001). As with science self-efficacy, gender differences in computer game self-efficacy may stem from a variety of social influences—home, education, cultural, mass media—that shape the experiences and perceptions that adolescents have with regard to gaming as a male stereotyped activity (Terlecki et al.). These messages and the different pattern of computer and video game use they engender for boys and girls may lead to the development of differential ability beliefs such as computer game self-efficacy (for a broader discussion of gender and self-efficacy, see Bussey and

Bandura 1999). Supporting this prediction, researchers have found that males have higher computer and video game self-efficacy than females (Ketelhut 2010; Terlecki et al.). For example, Ketelhut reported significantly higher scores on both computer and video gaming self-efficacy for middle school boys compared to girls. Gender differences in computer game self-efficacy could advantage boys in assessments based in IVEs, raising validity issues for this assessment format.

Current Study

The current investigation has two aims. The primary aim is to examine how achievement on a science assessment that requires enacting scientific inquiry skills in an IVE relates to prior and subsequent scientific inquiry self-efficacy beliefs. Prior research has been mixed on this point, and these relations have not been examined in the context of an IVE that is primarily an assessment activity. The current study adds to limited research in this area by examining whether scientific inquiry self-efficacy changes after a single session and as a function of achievement on the assessment. This aim addresses both the validity of the assessment as a test, in part, of scientific inquiry skill and the implications of such an assessment on subsequent motivational self-beliefs. To address the primary aim, we asked the following two research questions: (RQ1) To what extent does scientific inquiry self-efficacy explain variance in middle school students' performance in a challenge-based science assessment in an IVE? (RQ2) What effect does performance on this assessment have on subsequent science inquiry self-efficacy beliefs?

A second aim of the study is to evaluate the effect of the computer game format on performance. We do so to evaluate whether students who have low computer game self-efficacy may be unfairly disadvantaged in science assessments that take place via interactions in an IVE. In doing so, we add to a limited literature on the effect of computer game self-efficacy on performance in educational games and IVEs and, to our knowledge, are the first to examine this relationship for middle school science students. Given the historical trend for gender differences in computer game use and computer game self-efficacy beliefs, we evaluate whether boys and girls differ in their levels of computer game self-efficacy and whether these differences influence achievement in the assessment. In doing so, we evaluate gender differences on two types of self-efficacy beliefs with limited prior research: scientific inquiry self-efficacy and computer game self-efficacy. To address the secondary aim, we ask the following three research questions: (RQ3) To what extent does computer game self-efficacy explain variance in middle school students' performance in an IVE? (RQ4) What effect does performance on the assessment have on subsequent computer-game self-efficacy beliefs? (RQ5) Are there gender differences in scientific inquiry self-efficacy and computer game self-efficacy and their relations with performance on a science assessment in an IVE?

As described above, the current study examines the relations of scientific inquiry self-efficacy, computer game self-efficacy, and performance in a science assessment in an IVE. This research is part of a project called Situated Assessment using Virtual Environments for Science Content and Inquiry (SAVE Science). SAVE Science consists of a series of game-based assessment modules for middle school students (see Fig. 1). Applying classroom learning to solve a problem in the virtual world called *Scientopolis*, students interact with digital characters, gather data by using interactive virtual tools within the immersive world, develop hypotheses, and propose solutions to the assessment problems (see Materials section for more details).

Materials and Methods

Participants

Participants were 407 middle school students (51 % female) from sixth, seventh, and eighth grades (22, 75, and 4 %, respectively). Students were drawn from the science classrooms of 12 teachers in three urban and six near-urban schools (37 and 63 % of sample, respectively), which served ethnically and socioeconomically diverse communities. Specific information regarding students' race/ethnicity was not recorded due to a technical malfunction. Data were collected during the 2010–2011 and 2011–2012 academic years.

Fig. 1 A SAVE Science world: Sheep Trouble module

Materials

Scientific inquiry self-efficacy was measured with the previously validated SETS self-efficacy subscale (Ketelhut 2010). The measure consists of 12 items using a 5-point Likert scale (5 = Strongly agree; 1 = Strongly disagree) that tap students' efficacy beliefs for successfully carrying out tasks associated with scientific inquiry [e.g., *When I do an experiment, it is hard for me to figure out how data I collected answers the question* (reverse scored)]. The measure had good reliability before and after participating in the IVE (Cronbach's $\alpha = .85$ and $\alpha = .89$, respectively). Mean scores were used in analyses.

Computer game self-efficacy was measured with the previously validated SETS computer game self-efficacy subscale (Ketelhut 2010), which consists of five items that tapped students' efficacy beliefs for succeeding at computer games (e.g., *I can learn how to play any computer game if I don't give up*). The measure had adequate to good reliability before and after participating in the IVE (Cronbach's $\alpha = .76$ and $\alpha = .83$, respectively). Mean scores were used in analyses.

Immersive Virtual Environment (IVE). The IVE module was designed to assess both students' ability to conduct inquiry and their understanding of evolution topics of adaptation and speciation. In the virtual world, students were tasked with finding out why some sheep on Farmer Brown's farm were dying. By interacting with characters and objects, players learned that some sheep had recently been imported to the hilly farm from an island where the terrain was flat while others were domestic to the farm. Students could move around the farm and gather information and data about different sheep, which could then be analyzed using a set of in-world tools. In order to fully answer the question of why some sheep were dying, students needed to realize that newly imported sheep had different physical features (i.e., longer legs) that were maladaptive for reaching hilltops of their new rocky environment, where the lush grass was growing, thus disposing the imported sheep to grow hungry, weak, and sick on the flat areas of the farm where grass was sparse. Drawing this conclusion required that students use scientific inquiry skills to collect and synthesize multiple sources of information (e.g., measurements from healthy and sick sheep at different locations) and prior knowledge (biological adaptation concepts). Below we describe what students see and do in the module.

After choosing an avatar, the story-based scientific problem appears in a dialogue box: The farmer has recently added new sheep to his flock; the sheep were healthy when they arrived, but are now sick; townspeople believe that the sheep are cursed and should be killed; the farmer wants to know if there is a scientific explanation for the sick sheep which might save them. After learning about the problem, students are free to move about a medieval farm (see Fig. 1 for a screenshot). The terrain of the farm contains a grassy hill and a flat area with little grass where the barn and farmhouses are located. Two physically distinguishable types of sheep graze as they move about the farm: one type is rotund with short legs; a second type is skinny with long legs. When students collide with a sheep, they can use virtual tools to measure and record data on five variables: leg length, ear length, body length, gender, and age. By using a graphing tool, students can examine aggregated data for these variables for new, old, or both new and old sheep (e.g., comparing average age of new vs. old sheep). When students collide with a person (farmer or resident), students can choose questions to ask and read corresponding dialogue about a character's perception of the problem. When students collide with the farmer, they may indicate whether they have solved the mystery or wish to continue exploring the farm. When students believe they have solved the mystery, they answer a series of questions regarding their understanding of the problem.

The *end-of-module assessment* consisted of six multiple choice questions embedded in the virtual environment. All questions assessed the ability to apply scientific knowledge about adaptation and speciation to the context presented in the IVE; questions focused on sheep characteristics and their relations to health and adaptation on the farm [e.g., *What variable could you change on the new sheep to help them: Make the legs (a) shorter; (b) longer; (c) doesn't make a difference*].

Procedures

After obtaining informed consent, students completed an introductory practice module in which they were introduced to solving a narrative-based problem in the IVE, and practiced maneuvering an avatar, interacting with objects (e.g., talking with characters, measuring objects), and answering practice end-of-module assessment questions. Then, on a separate day and under the supervision of their science teacher, students completed an electronic questionnaire that assessed scientific inquiry self-efficacy and computer game self-efficacy (along with other motivational constructs and individual background information, not analyzed here, which were relevant to the larger study). Within four instructional days of taking the pre-module questionnaire, teachers supervised a class of students as they individually completed the assessment module under test-like conditions (e.g., no student talking). Students completed the modules in approximately 20–30 min. Immediately following the completion of the module, students completed an electronic post-module questionnaire that assessed scientific inquiry self-efficacy and computer game self-efficacy. Students received feedback on their performance in the module after the post-module questionnaire was completed. The median number of weeks that passed between the initial classroom instruction on relevant topics and the administration of the IVE assessment was 2 weeks.

Analytic Approach

Preliminary Analyses

Data were screened for accuracy of input, linearity, normality, multicollinearity, and multivariate outliers. Screening indicated no violations of assumptions for multivariate linear modeling or for *t* tests. One multivariate outlier was identified and removed. Intra-class coefficients (ICCs) were calculated for each dependent variable using teacher as the clustering variable; all ICCs were <.05, indicating the amount of nestedness within teacher was small and multilevel model was unnecessary (Bickel 2007). For path analysis, we used full information maximum likelihood estimator in *Mplus* 7.0 (Muthen and Muthen 1998–Muthén and Muthén 2012) to handle missing data.

Path Analyses

Path analyses facilitate the comparison of multiple relations within a single analytic model, thereby allowing the examination of the unique association between multiple sets of variables as well as indirect effects. As illustrated in Fig. 2, we tested a path model in which pre-module scientific inquiry self-efficacy and computer game self-efficacy predict module performance, which in turn predicts post-module scientific inquiry self-efficacy and computer game self-efficacy. This model tested two theorized relations. First, the model tests the extent to which scientific inquiry self-efficacy and computer game self-efficacy uniquely predict performance on a science assessment in a virtual environment. Second, the model tests the extent to which performance on the assessment explains variance in post-module self-efficacy beliefs after accounting for the effect of pre-module self-efficacy beliefs. Pre-module science inquiry self-efficacy and pre-module computer game self-efficacy are allowed to be correlated since task-specific efficacy beliefs can be related to general efficacy beliefs (Woodruff and Cashman 1993).

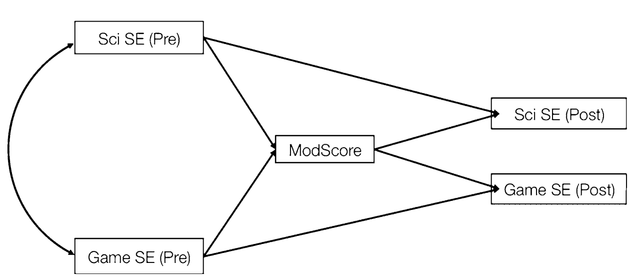

Fig. 2 Path model of scientific inquiry self-efficacy and computer game self-efficacy as predictors and outcomes of achievement in a science IVE. *Note* Sci SE = scientific inquiry self-efficacy; Game SE = computer game self-efficacy; ModScore = achievement in the virtual environment; Pre = before the assessment; Post = after the assessment

Multi-Group Path Model

A multi-group path analysis affords the opportunity to test whether model parameters are invariant across groups—that is, whether the model for different groups has substantively different relations. In the present study, differences in path coefficients by gender were of primary interest. To evaluate the extent to which path coefficients differed for males and females, we compared the fit indices of nested models, starting with the most restrictive model (all path coefficients were constrained to be equal for males and females). In a second step, a model was tested in which a single path coefficient was allowed to differ for males and females while all other paths were constrained to be equal across groups. The difference in Chi-square model fit for the more parsimonious model (all paths constrained to be equal) and the less parsimonious model (all but one path constrained to be equal) was tested for significance. Step two was repeated for each path in the model. Paths that resulted in a significantly improved model fit, if found, would be retained in the final model.

All results were analyzed at an alpha level of $p < .05$.

Results

Descriptive Statistics

Descriptive statistics and bivariate correlations for all variables in the path model by gender are presented in Table 1. All correlations were in the expected directions. As expected, pre- and post-module scientific inquiry self-efficacy scores were highly correlated with each other, as were pre- and post-module computer game self-efficacy scores. Pre- and post-module scientific inquiry self-efficacy scores tended to be significantly correlated with module scores, whereas pre- and post-module computer game self-efficacy scores tended not to be significantly correlated. As can be seen from the overall means for each variable, students tended to report moderate endorsement of scientific inquiry self-efficacy and computer game self-efficacy items. Students correctly answered just over half of the end-of-module questions, indicating that, on average, students had an incomplete understanding of the scientific

Table 1 Descriptive statistics and correlations for females (above diagonal) and males (below diagonal)

	1	2	3	4	5	Descriptives for females ($n = 181$)	
						M	*SD*
1. SciSE (pre)	–	.88	.42	.39	.24	3.50	0.61
2. SciSE (post)	.81	–	.37	.42	.32	3.49	0.62
3. GameSE (pre)	.18	.15	–	.75	.02	3.75	0.63
4. GameSE (post)	.13	.16	.80	–	.10	3.71	0.70
5. ModScore	.18	.17	.09	.26	–	0.55	0.25
Descriptives for males (n = 191)							
M	3.48	3.49	4.01	4.03	0.55		
SD	0.58	0.66	0.74	0.75	0.27		

Correlations for girls shown above the diagonal; correlations for boys shown below the diagonal; correlations >.17 are significant at $p < .05$; T1 = Before completing the module; T2 = After completing the module; Sci SE = scientific inquiry self-efficacy; Game SE = computer game self-efficacy; ModScore = percentage correct on end-of-module assessment.

problem presented in the module (i.e., why some sheep were sick).

Gender Differences in Mean Efficacy and Module Scores

In order to evaluate whether there were gender differences in scientific inquiry self-efficacy and computer game self-efficacy (RQ5), we conducted a series of independent samples t tests. Males had higher mean computer game self-efficacy scores pre-module ($t[364.879] = 3.53$, $p < .001$, $d = 0.53$) and post-module ($t[224] = 3.314$, $p < .001$, $d = 0.61$) than females. There were no statistically significant differences by gender between mean module scores ($t[260] = 1.190$, $p = .850$) or for scientific inquiry self-efficacy pre-module ($t[369] = -0.191$, $p = .858$) or post-module ($t[369] = -0.063$, $p = .950$).

Model Fitting

In order to evaluate the effects of pretest self-efficacy beliefs on achievement in the IVE (RQ1, RQ3) and the effects of achievement on subsequent self-efficacy beliefs (RQ2, RQ4), we tested the fit of Model 1. Model 1, which modeled the proposed relations on the entire sample regardless of gender, demonstrated good fit by conventionally accepted criteria (e.g., Kline 2010), indicating a close fit between the actual data and modeled relations. The model fit statistics for Model 1 are presented in the top portion of Table 2.

In order to evaluate whether path coefficients differed significantly by gender (RQ5; see Method section for more detail), we tested a series of models in which a single path coefficient was allowed to vary by gender while all other paths were constrained to be invariant. Model fit statistics for selected multi-group models are presented in Table 2. In all such models, the difference in model fit between the less and more parsimonious model was not statistically significant. Thus, the results from model fitting suggest that a model that does not estimate parameters to be different by gender was more parsimonious and fit the data well. We describe and discuss model parameters for the most parsimonious model (Model 1) only but for the purposes of comparison, we present path coefficients for Model 1 (whole sample) and Model 2d (no constraints for path coefficients by gender) in Table 3.

Model parameters for Model 1 are presented in Fig. 3 and Table 3. Pre-module scientific inquiry self-efficacy has a positive and statistically significant association with module score ($\beta = .221$), indicating a small effect. Higher scientific inquiry self-efficacy was associated with better module performance after accounting for the effect of computer game self-efficacy. The path between pre-module computer game self-efficacy and module scores was not statistically significant (RQ3). Thus, prior computer game self-efficacy did not predict module performance after accounting for the effect of prior scientific inquiry self-efficacy. Module scores were positively and statistically significantly associated with post-module scientific inquiry self-efficacy ($\beta = .170$) after accounting for initial scientific inquiry self-efficacy scores, indicating that variance in module score accounted for a small amount of change in scientific inquiry self-efficacy over time. Thus, students who earned higher module scores tended to report an increase in scientific inquiry self-efficacy. Conversely, students with lower module scores tended to report a decrease in scientific inquiry self-efficacy. Module score was not statistically significantly associated with post-module computer game self-efficacy (RQ4). Thus, for scientific inquiry self-efficacy—but not computer game self-efficacy—performance on the assessment predicted

Table 2 Summary of model fit for total group and two-group tests of invariance

Model	χ^2	*df*	*p*	RMSEA	CFI	TLI	SRMR	Model description
Total group analysis (*n* = 407)								
1	0.084	2	.959	<.001	1.000	1.016	0.003	Total sample
Two-group (male/female) analysis (*n* = 372)								
2a	0.684	4	.953	<.001	1.000	1.029	0.011	Total invariance (inv)
2b	1.877	7	.871	<.001	1.000	1.026	0.015	Sci SE inv
2c	2.069	7	.956	<.001	1.000	1.025	0.020	Game SE inv
2d	3.032	10	.981	<.001	1.000	1.024	0.021	Total invariance

All possible nested models between total invariance between males and females (Model 2a) and no invariance between males and females (Model 2d) by allowing a single path to differ by gender and evaluating the change in Chi-square model fit. For the purpose of demonstration, we present two nested models between total invariance (Model 2a) and no invariance (Model 2d) in which paths between ModScore and GameSE (Model 2b) or ModScores and SciSE (Model 2c) are allowed to vary by gender

SciSE, scientific inquiry self-efficacy; GameSE, computer game self-efficacy

Table 3 Maximum likelihood parameter estimates for total sample and multigroup path models

Parameter	Model 1			Model 2d					
	St.	*SE*	*p*	Boys			Girls		
				St.	*SE*	*p*	St.	*SE*	*p*
Direct effects									
SciSE(Pre) → ModScore	.221	.058	<.001	.167	.087	.053	.274	.088	.002
GameSE (Pre) → ModScore	.013	.060	.833	.054	.084	.522	−.083	.091	.363
ModScore → SciSE (post)	.170	.051	.001	.211	.073	.004	.113	.067	.092
ModScore → GameSE (post)	.104	.055	.059	.091	.076	.235	.145	.088	.097
SciSE (Pre) → SciSE (post)	.795	.023	<.001	.769	.035	<.001	.851	.029	<.001
GameSE (Pre) → GameSE (post)	.788	.023	<.001	.800	.031	<.001	.753	.041	<.001
Indirect effects									
SciSE (Pre) → ModScore → SciSE(post)	.038	.015	.010	.039	.024	.103	.031	.021	.129
Correlations									
SciSE (pre) with GameSE (pre)	.286	.046	<.001	.184	.070	.009	.418	.061	<.001
Error correlations									
SciSE (post) with GameSE (post)	.188	.068	.006	.205	.094	.030	.190	.100	.059
R^2									
SciSE (post)	.723	.031	<.001	.694	.048	<.001	.784	.036	<.001
GameSE (post)	.644	.035	<.001	.661	.047	<.001	.595	.062	<.001
ModScore	.051	.025	.041	.034	.031	.268	.063	.040	.117

St. = standardized coefficients; SciSE = scientific inquiry self-efficacy; GameSE = computer game self-efficacy; ModScore = end-of-module assessment score

subsequent efficacy beliefs beyond what was accounted for by initial beliefs. Scientific inquiry self-efficacy and computer game self-efficacy scores were themselves significantly correlated before the module ($\beta = .286$), as were the error terms for these variables after completing the module ($\beta = .188$). Results indicate a small but statistically significant indirect effect of pre-module scientific inquiry self-efficacy on post-module scientific inquiry self-efficacy via module scores ($\beta = .038$). The model accounted for statistically significant amounts of variance in post-module scientific inquiry self-efficacy (72.3 %), post-module computer game self-efficacy (64.4 %), and module scores (5.1 %).

Discussion

This study examined middle school students' performance on a science assessment in a computer game-like immersive virtual environment as it related to two types of self-efficacy beliefs: scientific inquiry self-efficacy and computer game self-efficacy. The primary aim of the study was

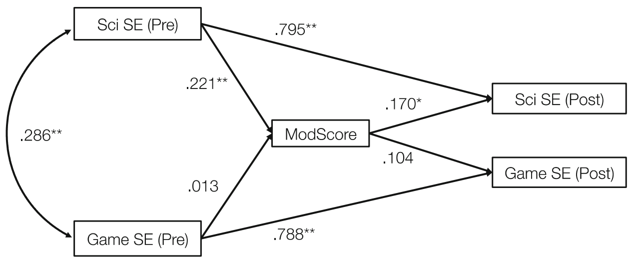

Fig. 3 Path model of scientific inquiry self-efficacy and computer game self-efficacy as predictors and outcomes of achievement in a science IVE. *Note* Sci SE = scientific inquiry self-efficacy; Game SE = computer game self-efficacy; ModScore = achievement in the virtual environment; Pre = before the assessment; Post = after the assessment. *$p < .01$ **$p < .001$

to examine how students' prior and subsequent scientific inquiry self-efficacy beliefs related to performance on a science assessment in an IVE. This aim was addressed through research questions examining the extent to which initial scientific inquiry self-efficacy explained variance in students' performance in an IVE (RQ1) and the effect this performance had on subsequent science inquiry self-efficacy beliefs (RQ2). We found that initial scientific inquiry self-efficacy significantly predicted achievement in the IVE; achievement, in turn, significantly predicted changes in scientific inquiry self-efficacy. As we describe below, these relationships document how assessment can shape subsequent self-efficacy beliefs and provide validity evidence for the assessment as a test of scientific inquiry skills. A secondary aim of the study was to examine whether achievement on the computer game-like assessment was equitable for students with different levels of computer game self-efficacy. This aim was addressed through research questions examining the extent to which initial computer game self-efficacy explained variance in performance in an IVE (RQ3), the effect of performance on subsequent computer-game self-efficacy beliefs (RQ4); and whether there were gender differences in scientific inquiry self-efficacy and computer game self-efficacy, and their relations with performance on a science assessment in an IVE. We found that computer game self-efficacy did not significantly predict performance in the IVE nor did performance significantly predict changes in subsequent computer game self-efficacy. Boys and girls had similar levels of scientific inquiry self-efficacy at both time points. We observed gender differences in mean computer game self-efficacy scores (males > females), but this difference did not appear to disadvantage girls on the assessment. As we describe below, these relationships document that assessments in IVEs can be equitable for boys and girls alike, despite gender differences in computer game self-efficacy beliefs.

We found that students' initial scientific inquiry self-efficacy predicted performance in an IVE that assessed science topic knowledge and inquiry skills, a finding that is consistent with prior research (e.g., Ketelhut 2007; Nelson and Ketelhut 2008). The current study extends these findings by documenting the relationship with a larger sample and across a more diverse set of schools (e.g., both urban and near-urban). The model accounted for only a modest amount of variance (5 %) in student performance in the module, indicating the great majority of variance in module performance was captured by variables not included the model (e.g., prior knowledge). The small amount of explained variance in performance is particularly of interest because prior meta-analysis (Multon et al. 1991) showed a mean of 14 % of the variance in performance explained by self-efficacy beliefs. One possible explanation is that self-efficacy may have less influence on performance in the context of a group-administered assessment, where the decision to persist with the task may be influenced by other factors, such as when peers finish the task and time constraints in the classroom. Future research might explore whether initial self-efficacy levels have less of an influence on performance and persistence in virtual environments than in other more traditional classroom contexts.

Our results extend prior literature by documenting that individual achievement in the IVE predicted change in scientific inquiry self-efficacy beliefs. Consistent with self-efficacy theory (Bandura 1986, 1997), mastery experiences in the module were positively associated with subsequent efficacy beliefs. Presumably, when students perceived success with the inquiry task, they tended to increase their self-efficacy for doing science, and vice versa.

Evidence of change in scientific inquiry self-efficacy as a function of success on the assessment tasks in the IVE provides evidence for validity of the assessment as a measure of scientific inquiry. Our findings differ from prior research. Our results showed increased scientific inquiry self-efficacy only for those students with high module scores and decreases in scientific inquiry self-efficacy for those with low module scores, resulting in similar group mean scientific inquiry scores before and after the module. By contrast, prior studies have found group-level increases in scientific inquiry self-efficacy (Chen et al. 2014) or science self-efficacy (Meluso et al. 2012) after participating in IVE-based activities, regardless of an individual's level of performance.

One possible explanation for this difference is the nature of the educational tasks. Whereas Chen et al. and Meluso et al. examined environments (EcoMUVE and Crystal Island, respectively) designed to facilitate *learning* of science content and inquiry skills, SAVE Science modules were designed to *assess* science knowledge and skill. Learning-centered designs may differ from assessment-centered designs in several ways; these include, among others, the amount of time to reach mastery (unlimited vs.

constrained), the presence and type of feedback (formative vs. summative), the availability and type of assistance (e.g., tips vs. no assistance), and the extent to which students interact with others (e.g., collaboration vs. independent work). Presumably, virtual environments designed to facilitate learning may have resulted in mastery for more students, and therefore, group mean increases in science-related self-efficacy beliefs were observed. By contrast, virtual environments that focus on assessing independent knowledge and ability with time constraints and without assistance, social interaction, and formative feedback are likely to result in greater variability in the extent to which students attain mastery; this may result in greater variability in the extent and direction of change in self-efficacy beliefs. Nonetheless, this shows the important impact of mastery experiences in an assessment and underscores the necessity to be sure that all assessments are valid and reliable across the domain and student group.

Regarding our secondary aim of evaluating the effect of the computer game-based format, our results indicate that initial computer game self-efficacy did not significantly predict performance in the module. This finding indicates that this particular assessment in a virtual environment was equitable no matter how students in this sample perceived their computer game playing abilities. While this study does not examine why initial computer game self-efficacy is unrelated to module performance, we speculate that several characteristics of the module make it substantively different than many computer or video games that use virtual environments. In the current study, the module emphasized an intellectual rather than a physical quest and restricted actions to exploring the terrain, talking with characters, and collecting data (e.g., taking measurements). Previous studies that found positive effects of video game self-efficacy and prior video game experience on performance were conducted using instructional games in which players completed tasks that more closely resemble commercial video games, such as marksmanship tasks (e.g., Orvis et al. 2008; Pavlas et al. 2010). Gaming abilities and computer game self-efficacy may be associated with achievement in computer game-based educational tasks when the tasks make strong computer gaming demands on the player. This is analogous to the significant relationship reading abilities have with achievement in reading-intensive science assessments (e.g., O'Reilly and McNamara 2007). A productive area of future research may be to examine the characteristics of virtual environments and the conditions and characteristics that make assessments equitable for both avid and beginner gamers. This has potential impact for educational game designers.

It is interesting to note that the correlation between computer game self-efficacy and module performance is stronger after completing the module compared to before completing the module. One explanation is that the term *computer games* refers to a family of quite diverse activities (e.g., solitaire vs. role playing). Students may have considered different types of computer games when reporting initial computer game efficacy beliefs. After completing the assessment in a virtual environment, student may have reported efficacy beliefs for succeeding at computer games that more closely resemble the module. To increase the validity of their measures, researchers might specify the type of games they wish respondents to consider when self-reporting game-related efficacy beliefs.

Turning to gender differences, we found significant differences in boys' and girls' computer game self-efficacy, consistent with Ketelhut (2010). Boys reported higher computer game self-efficacy scores both before and after the module compared to girls. This finding is consistent with the literature that indicates that, compared to girls, boys are more interested in computer games and play them more frequently (e.g., Lucas and Sherry 2004; Wright et al. 2001). Our results add to this literature by further documenting gender differences in self-efficacy beliefs for playing computer games. Gender differences in computer game self-efficacy likely result from differences in exposure to and experience with computer games as well as the gendered cultural messages boys and girls receive about computer game playing (Bussey and Bandura 1999; Terlecki et al. 2011). Despite gender differences in mean computer game self-efficacy, results from the multi-group modeling suggested only minor differences between boys and girls in the relations among efficacy beliefs and module performance, consistent with recent work by Nietfeld et al. 2014. The absence of significant gender differences in how self-efficacy beliefs relate to IVE performance is important as it suggests that the SAVE Science assessment is equitable with regard to gender, even while gender differences in computer game self-efficacy are observed.

Importantly, girls and boys did not differ in their scientific inquiry self-efficacy scores before or after completing the module. Historically, research on gender differences in science self-efficacy beliefs in middle school students has been mixed, but our results join a trend of studies finding no gender differences (Britner and Pajares 2001, 2006; Chen and Usher 2013; Kiran and Sungur 2012; Usher and Pajares 2008). The current study adds to this literature by examining self-efficacy for engaging in scientific inquiry tasks specifically rather than being efficacious in science class more broadly. In doing so, we document comparable levels of scientific inquiry self-efficacy among boys and girls in middle school.

Limitations and Implications

Interpretation of the results should be considered in parallel with the following limitations. According to Bandura (1986, 1997), it is the *interpretation* of experiences—not the experiences per se—that influences academic self-efficacy beliefs. In the current study, we did not measure students' *perceived* level of success in the IVE or their attributions for perceived success or failure experiences. Given that actual performance in the module predicted change in scientific inquiry self-efficacy, our findings suggested that actual and perceived success were aligned. Nevertheless, future research might directly measure students' perceptions about performance experiences (e.g., level of success and attributions).

Results from this study have implications for practice, IVE design, and future research. First, given the differences in middle school boys' and girls' computer game self-efficacy beliefs, educators should be aware that boys and girls may approach computer game educational tasks with different ability beliefs, which may shape their willingness to engage with or persist at such tasks. Given this gender difference, we advise that researchers measure computer game self-efficacy to evaluate whether such beliefs influence engagement and achievement in computer game-like IVEs. Second, we demonstrate that, despite gender differences in computer game self-efficacy, science assessments in IVEs can be equitable with regard to gender and computer game self-efficacy. That is, gender differences in computer gaming should not rule out the use of computer game-like IVEs for educational purposes, including for assessing scientific skill and ability. Third, science assessments in IVEs can make measurable differences in students' beliefs in their ability to succeed at scientific inquiry. This highlights the potential of scientific inquiry activities in IVEs to strengthen or weaken scientific inquiry self-efficacy. Providing feedback, strategy training, and multiple opportunities to achieve mastery may help guard against weakening self-efficacy beliefs, which may result for students who fail to achieve initial success.

In the current study, computer game self-efficacy was not predictive of or predicted by performance in a virtual environment. The nonsignificant relations among computer game self-efficacy and performance in the IVE presumably reflected the particular nature of this virtual environment and the extent to which it made computer game-like demands on the player. Our findings suggest that the challenge of the SAVE Science module examined in the current study was primarily a scientific inquiry challenge rather than a computer game challenge. This is important because one draw of educational video and computer games, including IVEs, is that their game-like nature makes educational content more appealing (Lepper and Malone 1987; for a recent discussion, see Habgood and Ainsworth 2011). Thus, to ensure that assessment instruments that are similar to computer or video games are equitable for students with different levels of gaming experience, it is important to evaluate the extent to which student performance in such settings is a function of gaming skills. The design features of virtual environments determine the nature of the experience and thus, presumably, the extent to which computer game skills and competency beliefs may influence performance. We recommend that researchers and designers of virtual environments with educational purposes consider the extent to which such environments provide equitable access for students who differ in their experience and perceptions as computer game players.

Conclusion

Our findings indicate that students change their scientific inquiry self-efficacy beliefs in relation to their achievement in a science assessment in an IVE. Our results provide evidence for the concurrent validity of SAVE Science modules as an assessment of science knowledge and inquiry skill rather than a test of gaming ability. Further, the results indicate that the assessment is equitable by gender, despite gender differences in computer gaming perceptions. Results suggest that students perceived the challenges in the IVE as authentic assessments of their ability to be successful at science inquiry tasks, lending weight to the potential of IVEs to function as fair assessments of scientific inquiry skills for a broad range of students.

Acknowledgments We acknowledge the contributions of Brian Nelson, Catherine Schifter, Martha Caray, Mandy Kirchgessner, Chris Teufel, Angela Shelton, and other SAVE Science team members for their contributions to the larger project of which this study is a part. We also are grateful to Jennifer Cromley and anonymous reviewers for their helpful feedback on previous versions of this manuscript. This material is based upon work supported by the National Science Foundation under Grant No. 0822308.

Conflict of interest The authors declare that they have no conflicts of interest.

References

Bandura A (1977) Self-efficacy: towards a unifying theory of behavioural change. Psychol Rev 84:191–215
Bandura A (1986) Social foundations of thought and action: a social cognitive theory. Prentice Hall, Englewood Cliffs
Bandura A (1997) Self-efficacy: the exercise of control. Freeman, New York
Bandura A, Barbaranelli C, Caprara GV, Pastorelli C (2001) Self-efficacy beliefs as shapers of children's aspirations and career trajectories. Child Dev 72(1):187–206. doi:10.1111/1467-8624.00273

Bickel R (2007) Multilevel analysis for applied research: it's just regression!. Guilford Press, New York

Britner SL, Pajares F (2001) Self-efficacy beliefs, motivation, race, and gender in middle school science. J Women Minor Sci Eng 7:271–285

Britner SL, Pajares F (2006) Sources of science self-efficacy beliefs of middle school students. J Res Sci Teach 43(5):485–499. doi:10.1002/tea.20131

Buchman DD, Funk JB (1996) Video and computer games in the '90 s: children's time commitment and game preference. Child Today 24:12–15

Bussey K, Bandura A (1999) Social cognitive theory of gender development and differentiation. Psychol Rev 106(4):676

Chen JA, Usher EL (2013) Profiles of the sources of science self-efficacy. Learn Individ Differ 24:11–21. doi:10.1016/j.lindif.2012.11.002

Chen JA, Metcalf SJ, Tutwiler MS (2014) Motivation and beliefs about the nature of scientific knowledge within an immersive virtual ecosystems environment. Contemp Educ Psychol 39:112–123

Clark DB, Nelson B, Sengupta P, D'Angelo CM (2009) Rethinking science learning through digital games and simulations: genres, examples, and evidence. Invited topic paper in the proceedings of the national academies board on science education workshop on learning science: computer games, simulations, and education, Washington

Clarke-Midura J, Dede C (2010) Assessment, technology, and change. J Res Technol Educ 42(3):309–328

Eccles JS, Adler T, Futterman R, Goff SB, Kaczala C, Meece JL et al (1983) Expectations, values, and academic behaviors. In: Spence JT (ed) Achievement and achievement motivation. Freeman, San Francisco, pp 75–146

Habgood MPJ, Ainsworth SE (2011) Motivating children to learn effectively: exploring the value of intrinsic integration in educational games. J Learn Sci 20(2):169–206. doi:10.1080/10508406.2010.508029

Hartmann T, Klimmt C (2006) Gender and computer games: exploring females' dislikes. J Comput Med Commun 11(4):910–931

Hill C, Corbett C, St. Rose A (2010) Why so few? Women in science, technology, engineering, and mathematics. American Association of University Women, Washington, DC

Johnson RD, Hornik S, Salas E (2008) An empirical examination of factors contributing to the creation of successful e-learning environments. Int J Hum Comput Stud 66(5):356–369

Ketelhut DJ (2007) The impact of student self-efficacy on scientific inquiry skills: an exploratory investigation in River City, a multi-user virtual environment. J Sci Educ Technol 16(1):99–111

Ketelhut DJ (2010) Assessing gaming, computer and scientific inquiry self-efficacy in a virtual environment. In: Annetta LA, Bronack S (eds) Serious educational game assessment: practical methods and models for educational games, simulations and virtual worlds. Sense Publishers, Amsterdam, pp 1–18

Ketelhut DJ, Nelson B, Schifter C, Kim Y (2013) Improving science assessments by situating them in a virtual environment. Educ Sci 3(2):172–192

Kıran D, Sungur S (2012) Middle school students' science self-efficacy and its sources: examination of gender difference. J Sci Educ Technol 21(5):619–630. doi:10.1007/s10956-011-9351-y

Klimt C, Hartmann T (2006) Effectance, self-efficacy, and the motivation to play video games. In: Vorderer P, Bryant J (eds) Playing video games: motives, responses, and consequences. Lawrence Erlbaum, Hillsdale, pp 153–168

Kline RB (2010) Principles and practice of structural equation modeling, 3rd edn. Guilford Press, New York

Lapan RT, Shaughnessy P, Boggs K (1996) Efficacy expectations and vocational interests as mediators between sex and choice of math/science college majors: a longitudinal study. J Vocat Behav 49(3):277–291

Lepper MR, Malone TW (1987) Intrinsic motivation and instructional effectiveness in computer-based education. In: Snow RE, Farr MJ (eds) Aptitude, learning and instruction: III. Conative and affective process analyses. Erlbaum, Hillsdale, pp 255–286

Lucas K, Sherry JL (2004) Sex differences in video game play: a communication-based explanation. Commun Res 31(5):499–523. doi:10.1177/0093650204267930

Meluso A, Zheng M, Spires HA, Lester J (2012) Enhancing 5th graders' science content knowledge and self-efficacy through game-based learning. Comput Educ 59(2):497–504

Mikropoulos TA, Natsis A (2011) Educational virtual environments: a ten-year review of empirical research (1999–2009). Comput Educ 56(3):769–780. doi:10.1016/j.compedu.2010.10.020

Multon K, Brown S, Lent R (1991) Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. J couns psychol 38(1):30–38

Muthén LK, Muthén BO (2012) Statistical analysis with latent variables. Mplus user's guide: 7th edition (Version 7.0), Los Angeles

National Research Council (2005) America's lab report: investigations in high school science. National Academies Press, Washington

Nelson BC, Erlandson BE (2012) Design for learning in virtual worlds. Interdisciplinary approaches to educational technology. Routledge, New York

Nelson BC, Ketelhut DJ (2008) Exploring embedded guidance and self-efficacy in educational multi-user virtual environments. Int J Comput Support Collab Learn 3(4):413–427

Nietfeld JL, Shores LR, Hoffmann KF (2014) Self-regulation and gender within a game-based learning environment. J Educ Psychol 106(4):961–973. doi:10.1037/a0037116

O'Reilly T, McNamara DS (2007) The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. Am Educ Res J 44(1):161–196. doi:10.3102/0002831206298171

Orvis KA, Horn DB, Belanich J (2008) The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. Comput Hum Behav 24(5):2415–2433

Pajares F (1996) Self-efficacy beliefs in academic settings. Rev Educ Res 66:543–578

Pajares F (2002) Gender and perceived self-efficacy in self-regulated learning. Theory Pract 41(2):116–125

Pavlas D, Heyne K, Bedwell W, Lazzara E, Salas E (2010) Game-based learning: the impact of flow state and videogame self-efficacy. In: Proceedings of the human factors and ergonomics society annual meeting Vol. 54, No. 28, SAGE Publications, Beverley Hills, 2398–2402

Schunk DH (1983) Ability versus effort attributional feedback: differential effects on self-efficacy and achievement. J Educ Psychol 75:848–856

Schunk DH, Meece JL (2005) Self-efficacy development in adolescence. In: Pajares F, Urdan T (eds) Self-efficacy beliefs of adolescents. Information Age, Charlotte, pp 71–96

Stajovic AD, Luthans F (1998) Self-efficacy and work-related performance: a meta analysis. Psychol Bull 124:240–261

Tai R, Liu C, Maltese A, Fan X (2006) Career choice: enhanced: planning early for careers in science. Science 312(5777):1143–1144

Terlecki M, Brown J, Harner-Steciw L, Irvin-Hannum J, Marchetto-Ryan N, Ruhl L, Wiggins J (2011) Sex differences and similarities in video game experience, preferences, and self-efficacy: implications for the gaming industry. Curr Psychol 30(1):22–33

Timms M, Clements D, Gobert J, Ketelhut DJ, Lester J, Reese D, Wiebe E (2012) New measurement paradigms. Report by the new measurement paradigms group prepared for the community for advancing discovery research in education. Retrieved from http://cadrek12.org/sites/default/files/NMP%20Report%20041412_0.pdf

Turkle S (1995) Life on the screen: identity in the age of the internet. Simon and Schuster, New York

Usher EL, Pajares F (2008) Sources of self-efficacy in school: critical review of the literature and future directions. Rev Educ Res 78(4):751–796. doi:10.3102/0034654308321456

Wang MNM, Wu KC, Iris Huang TC (2007) A study on the factors affecting biological concept learning of junior high school students. Int J Sci Educ 29(4):453–464

Weiner B (2005) Motivation from an attribution perspective and the social psychology of perceived competence. In: Elliot AJ, Dweck CS (eds) Handbook of Competence and Motivation. Guilford Press, New York, pp 73–84

Wigfield A, Eccles JS (2002) The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In: Wigfield A, Eccles JS (eds) Development of achievement motivation. Academic Press, San Diego, pp 91–120

Woodruff SL, Cashman JF (1993) Task, domain, and general efficacy: a reexamination of the self-efficacy scale. Psychol Rep 72:423–432

Wouters P, van Nimwegen C, van Oostendorp H, van der Spek ED (2013) A meta-analysis of the cognitive and motivational effects of serious games. J Educ Psychol 105(2):249

Wright JC, Huston AC, Vandewater EA, Bickham DS, Scantlin RM, Kotler JA, Finkelstein J (2001) American children's use of electronic media in 1997: a national survey. J Appl Dev Psychol 22(1):31–47. doi:10.1016/S0193-3973(00)00064-2

Zimmerman BJ (2000) Self-Efficacy: an essential motive to learn. Contemp Educ Psychol 25(1):82–91. doi:10.1006/ceps.1999.1016