Interactive Learning Environments, 2014
http://dx.doi.org/10.1080/10494820.2014.881393

Routledge
Taylor & Francis Group

Effects of an instructional gaming characteristic on learning effectiveness, efficiency, and engagement: using a storyline for teaching basic statistical skills

Elena Novak[a,b*], Tristan E. Johnson[a,c], Gershon Tenenbaum[a] and Valerie J. Shute[a]

[a]*Department of Educational Psychology and Learning Systems, Florida State University, Tallahassee, FL, USA;* [b]*School of Teacher Education, Western Kentucky University, 1906 College Heights Blvd. #61030, Bowling Green, KY 42101, USA;* [c]*Graduate School of Engineering, Northeastern University, 130 Snell Engineering Center; 360 Huntington Avenue, Boston, MA 02115, USA*

(*Received 3 June 2013; final version received 9 December 2013*)

The study explored instructional benefits of a storyline gaming characteristic (GC) on learning effectiveness, efficiency, and engagement with the use of an online instructional simulation for graduate students in an introductory statistics course. A storyline is a game-design element that connects scenes with the educational content. In order to examine the interactions between the storyline GC and human performance, a storyline was embedded in a simulation. The goal of the simulation was to engage students in problem-solving and data analysis in the context of basic statistics by using real-world examples. The authors developed two different versions of the simulation: (1) Simulation+No GC, and (2) Simulation+Storyline GC. Both versions shared the same instructional content but differed in the presence or absence of a storyline GC. The results indicated that adding a storyline to a simulation did not result in significant improvements in learning effectiveness, efficiency, or engagement. However, both instructional methods (simulation and simulation with a storyline) showed significant learning gains from pre- to post-test. The findings of this study offer future directions for embedding a storyline GC into learning content.

Keywords: gaming characteristics; storyline; simulations; games; statistics education

There has been a growing body of the literature examining instructional computer-based games and their effects on academic achievement. In addition to the traditional use of games in early childhood education, greater numbers of educators and employers are using games for teaching complex concepts in K-12 and higher education settings, as well as training employees in business, the industry, and the military. Game-based learning has captured educators' and researchers' attention as a means to enhance the effectiveness of learning by making the learning context more appealing, memorable, and engaging than traditional contexts (Gee, 2003).

There are many potential benefits to games. Research shows that a game-based learning approach can facilitate hands-on student-centered learning (case-based approach) (Watson, Mong, & Harris, 2011), encourage integration of knowledge from different areas to make a

*Corresponding author. Email: elena.novak@wku.edu; eno6c@my.fsu.edu

decision, and to examine the outcomes, e.g., Smithtown (Shute & Glaser, 1990), Taiga Park (design-based approach) (Barab et al., 2007). Games are believed to promote ownership of learning by providing the learner with control over the gaming process and difficulty levels (experimental approach) (Gee, 2008). Games can facilitate active learning or learning by doing that not only affects the learning outcomes, but keeps the learner engaged and motivated. Furthermore, games promote situated cognition where learners are immersed in an authentic learning environment during the game play (Squire, 2005).

The principles of designing good games that promote effective learning should be grounded in learning theories focusing on (a) gaming characteristics (GCs) that contribute to engagement, (b) architecture and functionality of computer-based games, and (c) assessment mechanisms embedded in the games (Gee, 2003). To apply learning principles for each of these components, there is a need for research aimed at developing a body of knowledge about specific factors and best practices to guide the game development process and design practices. For example, there is a very little knowledge about the effects of GCs on learning and engagement (Wilson et al., 2009).

While the body of knowledge is still developing, GCs provide a framework that explicates what game-design elements are critical for an instructional program to be considered a game (Federation of American Scientists, 2006; Wilson et al., 2009). However, there is little empirical evidence of linking GCs with performance. Exploring their effects on learning is a very complicated task since it requires understanding the role of a single GC in the learning process. In addition, it is important to explore the effects of multiple GCs combined, since most games usually include more than one GC (e.g., Shute & Ke, 2012).

The empirical research in the area of GCs typically centers on the effects of single GCs on human performance. The research design of studies exploring GCs usually involves embedding various GCs into a computer-based learning environment and comparing a GC environment with a non-GC one (DeRouin-Jessen, 2008; Parker & Lepper, 1992). Some researchers used the opposite approach of stripping down single GCs from an existing game (Malone, 1981). Among various GCs identified as critical for an instructional program to be considered a game are challenge, competition, rules, goals, storyline, changed reality, immediate feedback, and control over the pace of learning (Wilson et al., 2009).

The majority of empirical research investigating instructional benefits of GCs has focused so far on feedback (Shute, 2008), and control over the pace of learning (Azevedo, 2005). Several scientific endeavors explored the effects of a storyline/fantasy on learning and motivation (DeRouin-Jessen, 2008; Greenwood-Ericksen, 2008; Malone, 1981; Parker & Lepper, 1992).

Sid Meier, a developer of games such as *Pirates! Railroad Tycoon, Covert Action*, and *Civilization* described the role of storyline in games by comparing digital games with movies in his interview for Rouse (2005, p. 30):

> We try and put that amount of realism and accuracy into the game. And then make it fun on top of that. In the same way that a movie gives you all the fun and the action sequences and all the important parts of a story and then jumps quickly over the boring things, I think the game has the same responsibility, to bring you to the key decision points and then move you on to the next interesting thing.

Not surprisingly, storyline was identified among the major reasons people play video games (Green, Brock, & Kaufman, 2004). One of the major benefits associated with including a storyline into learning materials is students' increased interest and curiosity (Lee &

Chen, 2009; Malone, 1981). Moreover, claims have been made that a storyline embodies features of contextual anchoring and provides a cognitive framework (Dickey, 2006). On the other hand, following a storyline requires learners to pay attention to the storyline events and devote cognitive resources to process this information (McKee, 2005). One question still not answered is whether the cost of attending to the storyline outweighs the learning gains achieved through the increased contextual anchoring, interest, and curiosity (Ang, Zaphiris, & Mahmood, 2007).

Many games include some form of a storyline, fantasy, or narrative to create a conceptual framework of the learning context. In fact, embedding a story or narrative in learning content has become a growing trend in instructional gaming, since it allows educators to develop games or game-like environments that have educational value, e.g. the Writing Environment for Educational Video (Marchiori et al., 2012), and the Data-oriented Learning System of Statistics based on Analysis Scenario/Story (Mori, Yamamoto, & Yadohisa, 2003). However, research concerning the effects of a storyline on human learning and performance is very limited, and the findings are mixed. For instance, Parker and Lepper (1992) compared learning effectiveness of computer-based learning materials with fantasy/storyline "embellishments" vs. no-fantasy version. The results indicated that embedding fantasy elements into the learning content had a significant positive effect on children's motivation. Spires, Turner, Rowe, Mott, and Lester (2010) created three versions of a learning environment for testing instructional benefits of a storyline GC: full narrative, narrative-light, and no-narrative versions. The results indicated that students from the no-narrative group significantly outperformed students from the full narrative and narrative-light groups on the science content learning.

In order to further develop our understanding of the effects that storyline plays on learning, this study was designed with the purpose of measuring the instructional benefits of a storyline on learning effectiveness, efficiency, and engagement, i.e., e^3-learning, with the use of a participatory simulation that guided players to perform specific tasks within a given setting. Specifically, we used an approach of designing an instructional simulation and enhancing it with a storyline. Two different versions of the simulation were developed: (1) Simulation+No GC, and (2) Simulation+Storyline GC. Both versions had the same content, designed to teach basic introductory statistics concepts, but differed in the presence or absence of a storyline.

The effectiveness of the simulation-based learning was measured by using a post-test where students were asked to define and apply the learned statistical concepts and rules. Learning efficiency was measured by the ratio of students' post-test performance scores to the time spent on simulation-based activity. Students' engagement was assessed via Song and Keller's (2001) simplified version of Instructional Material Motivation Survey (IMMS) developed originally by Keller (1993).

Research hypothesis: Adding a storyline to a simulation will improve students' learning effectiveness, efficiency, and engagement.

In this study, we expected that adding a storyline to the learning simulation would improve students' learning effectiveness. This hypothesis was based on empirical findings showing some positive effects of including a storyline into learning content (Cordova & Lepper, 1996) as well as numerous theoretical claims in support of enhancing learning materials with a storyline as a means to motivate and engage students (Habgood, Ainsworth, & Benford, 2005). Given the proximity of statistics and mathematics disciplines, mathematics word problems research perspective can be possibly applied to storyline-enhanced problems, suggesting that placing statistics in context helps students transfer their statistics knowledge and skills into real-life situations (Walkington, Sherman, & Petrosino, 2012).

Adding a storyline to a learning simulation was also hypothesized to improve the learning efficiency of a learning simulation. This hypothesis was based on the assumption that adding a storyline to a simulation would not substantially increase the time spent on simulation-based activities, but would significantly improve students' learning. It was also expected that adding a storyline to a learning simulation would enhance students' engagement. Based on the findings from the literature regarding the instructional effectiveness of a storyline (Habgood et al., 2005), we expected that a storyline would "hook" students' attention thereby enhancing their engagement.

We conducted several small group and one-on-one formative evaluations with undergraduate and graduate students at various stages of the storyline and simulation development phases in order to understand which student audience would benefit the most from our learning materials. This continuous evaluation process revealed that undergraduate students did not find our storyline particularly appealing or relevant and that they had very low interest in learning statistics outside of the classroom. On the other hand, graduate students had a favorable attitude toward the storyline and expressed an interest in learning statistics by using our simulations.

Method

Participants

The participants for this study included 64 graduate students (age $M = 28$, SD = 6.55) with basic statistical knowledge (no more than an introductory course in Descriptive and Inferential Statistics) at a major Southeastern public university. Among them were 47 females and 17 males. The demographic make-up consisted of 18 Asian/Pacific Islander, 35 Caucasian/White, 5 Black, 2 Multiracial, 1 Hispanic, and 3 classified as "other." Education graduate students were recruited who either had taken one introductory statistics course or were enrolled in an introductory statistics course during the term, in which the study was conducted.

Tasks and materials

An online instructional introductory statistics module was developed in the Moodle Learning Management System to promote students' knowledge and skills gained in their previous coursework. Although the learning content of the instructional module was aligned with an introductory graduate statistic course, the module was not part of the course. The materials for this study consisted of five components: (1) overview of statistics concepts to-be-practised, (2) pre-test assessing students' prior knowledge of the subject matter, (3) simulation-based learning environment for practising the newly acquired knowledge, (4) post-test assessing the instructed skills and concepts, and (5) demographic and engagement surveys.

The instructional materials for this study were developed through a careful examination of four introductory statistics textbooks, and instructional materials from three introductory statistics courses. The concepts for this study included standard deviation, variance, variability, and the empirical rule (SDVVER). An instructional goal analysis was conducted in collaboration with two professors from the Statistics and Measurement Department to identify relevant skills and knowledge to be covered.

The goal of the statistics learning simulation was to engage students in problem solving, data analysis, and practice of their newly acquired skills through analysis of real-world cases. Both simulations created within the Moodle Learning Management System included the following two instructional features that were found beneficial to simulation-based learning (Chang, Chen, Lin, & Sung, 2008):

- Immediate formative feedback to inform learners why the action taken was or was not appropriate for solving a particular task. According to Shute (2008), the research in the area of feedback-enhanced instruction shows that formative feedback is a powerful tool for improving learning processes and outcomes.
- Control over the pace of learning. According to Azevedo (2005), learners should be assisted in regulating their learning process. Learners should be able to control and self-monitor their learning process.

Instructional methods

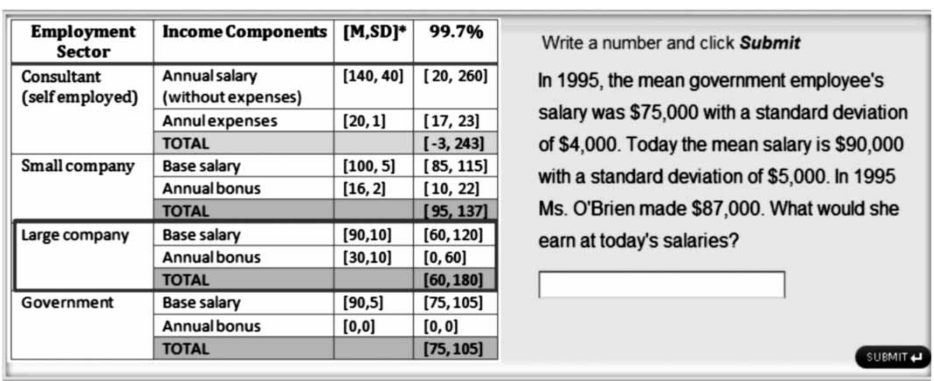
Two versions of the simulation were developed. Both versions shared the same content and instructional features but differed in the presence or absence of a storyline. The first version, Simulation+No GC, did not employ the storyline. The second version, Simulation+Storyline GC, integrated an overarching storyline into the learning materials to promote learner engagement.

Simulation+No GC instructional method

This instructional method included the most basic version of the simulation and served as a comparison for the storyline-enhanced simulation. The Simulation+No GC instructional method presented students with real-world problems requiring analysis of different employment options based on basic descriptive statistics data, such as a mean and standard deviation. Figure 1 illustrates an example of a problem from this instructional method.

Simulation+Storyline GC instructional method

The Simulation+Storyline GC instructional method presented learners with the Career Coach Simulation that had an embedded storyline. The Career Coach Simulation had the same learning content as the non-storyline instructional method. The storyline was developed by following the guidelines by Habgood et al. (2005) of how to create learning content intrinsically integrated into games.



Employment Sector	Income Components	[M,SD]*	99.7%
Consultant (self employed)	Annual salary (without expenses)	[140, 40]	[20, 260]
	Annul expenses	[20, 1]	[17, 23]
	TOTAL		[-3, 243]
Small company	Base salary	[100, 5]	[85, 115]
	Annual bonus	[16, 2]	[10, 22]
	TOTAL		[95, 137]
Large company	Base salary	[90,10]	[60, 120]
	Annual bonus	[30,10]	[0, 60]
	TOTAL		[60, 180]
Government	Base salary	[90,5]	[75, 105]
	Annual bonus	[0,0]	[0, 0]
	TOTAL		[75, 105]

Figure 1. Simulation+No GC intervention: practice problem preview.

Specifically, we used an in-game storytelling approach for developing the storyline-enhanced simulation. That is, the simulation unraveled the storyline details while students interacted with the simulation-based activities as opposed to presenting a story out of a game, when the storytelling occurs while players are not playing the game (Rouse, 2005). In this way, the learning content was integrally built into the storyline.

The Career Coach Simulation introduced a learner to a career-advising company setting, where the learner had to advise clients on how to find a "job of their dreams" by systematically examining various employment parameters such as income, job stability, amount of travel, flexibility, etc. Guided by the simulation storyline, the student adopted the role of a career coach. Through the interactions with a client, the career coach performs statistical analyses of various job parameters across four job sectors: government, small company, large company, and private business. The learning environment presented the learner with animations illustrating the storyline dialogs. Figure 2 shows a screenshot from the Career Coach Simulation.

Dependent variables

The dependent variables were as follows: effectiveness, efficiency, and engagement. Learning effectiveness was measured by a student's scores on the pre- and post-tests. Learning efficiency was considered the ratio of a student's post-test performance score and the time spent on using the simulation. Engagement was measured by using the simplified IMMS instrument (Song & Keller, 2001).

Effectiveness measure – the SDVVER instrument

The SDVVER learning assessment instrument was developed to assess how well students learned the skills and concepts introduced in the instructional module. Two parallel versions of the SDVVER instrument were developed, assessing skills specified under the learning

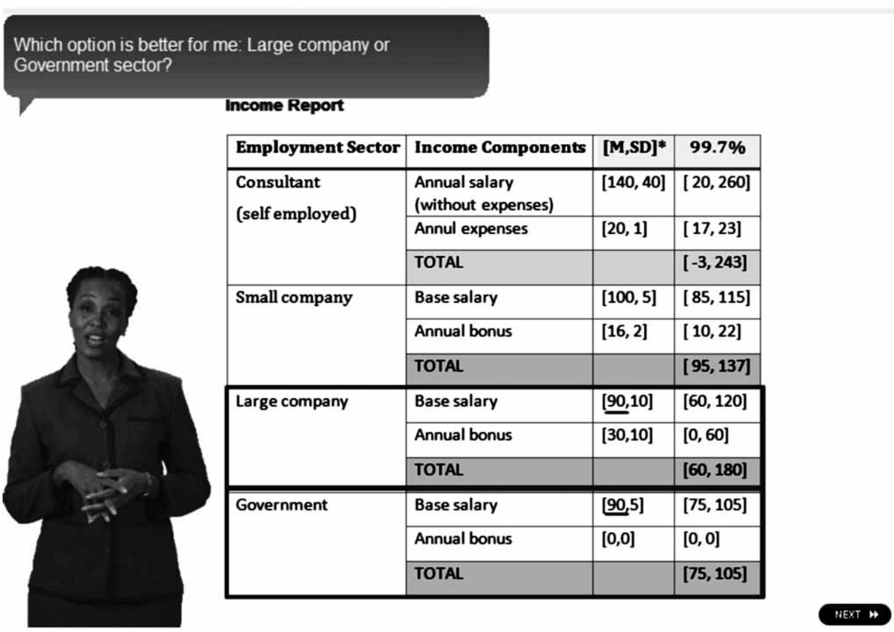
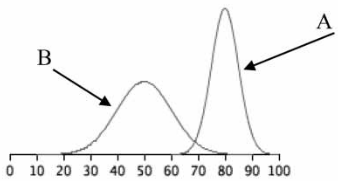
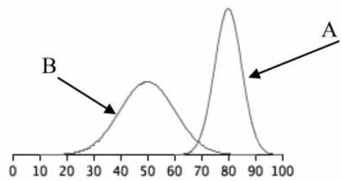


Figure 2. Simulation+Storyline GC intervention: Career Coach Simulation preview.

Table 1. The SDVVER instrument: parallel items assessing conceptual knowledge of standard deviation concept.

Version A	*Version B*
What distribution has the <u>larger</u> standard deviation?  • The A distribution • ✓ The B distribution	What distribution has the <u>smaller</u> standard deviation?  • ✓ The A distribution • The B distribution

objectives of the module that were identical for the two intervention conditions. For example, Table 1 presents two parallel items designed to assess conceptual knowledge related to the standard deviation concept. The SDVVER Version A was administered during the pre-test to measure whether a student had prior knowledge of the skills and concepts that are to-be-learned during the instructional module. The SDVVER Version B was administered during the post-test to measure how well students learned the instructional content.

Each version of the SDVVER instrument consisted of 18 items: 14 multiple-choice questions, 2 short numerical open questions, and 1 case-based problem that included 6 short numerical questions.

Engagement measure – the IMMS instrument (Song & Keller, 2001)

A simplified version of IMMS measures students' attitudes and motivation toward the instructional materials by looking at attention, relevance, confidence, and satisfaction (ARCS). Keller's (2008) analysis of ARCS categories in technology-supported learning systems demonstrated that engagement is essentially addressed in all four ARCS categories. The simplified IMMS instrument was validated in a number of studies. The simplified IMMS (Song & Keller, 2001) instrument included 12 items, where each of the 4 ARCS components was operationalized by 3 items. Students responded on a 5-point Likert-type scale with response choices: 1 (not true), 2 (slightly true), 3 (moderately true), 4 (mostly true), and 5 (very true). The potential total IMMS score ranges between 12 and 60. According to Song and Keller (2001), Cronbach's internal consistency was 0.92 for the overall IMMS score, 0.79 for attention, 0.73 for relevance, 0.70 for confidence, and 0.85 for satisfaction. Preliminary analysis of Keller's IMMS indicated some situational validity by using both exploratory and confirmatory factor analyses procedures (Huang, Huang, Diefes-Dux, & Imbrie, 2006).

Procedure

Participants were randomly assigned to one of the two intervention groups. The first author met in person with small groups of participants in a computer lab and provided individual

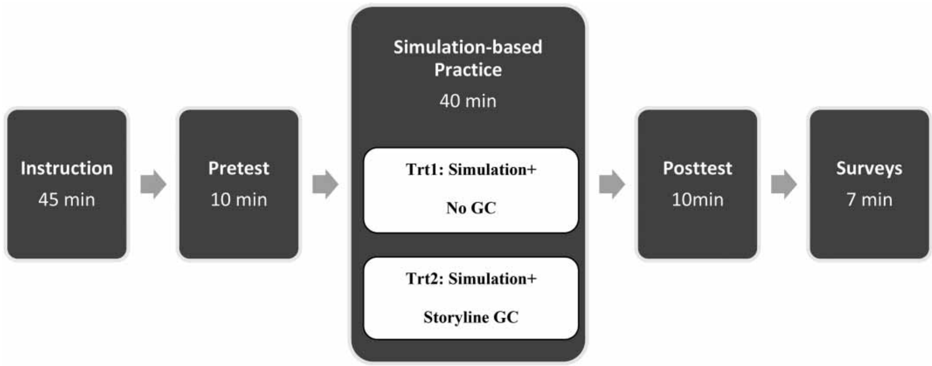


Figure 3. An overview of the study's procedure.

guidance to each participant on how to complete the experiment activities. The participants were informed neither about the real goal nor the research design of the study. There were no interactions among the participants during the experiment. Both intervention groups completed an online instructional module that required between 90 and 120 minutes per learner. During the first part of the module, students received the same online instruction that was designed to teach standard deviation and empirical rule concepts, which took approximately 45 minutes to complete for a student without a substantial background in the subject matter. Since some of the participants were taking an introductory statistics course during the same term in which the current study was conducted, and some of the participants took a similar course during their prior coursework, the goal of the online instruction was to review concepts-to-be practiced in the simulation. Therefore, in order to evaluate the effects of the simulation-based practice on e^3-learning, participants completed the pre-test immediately after the instructional part of the module. Following the pre-test, students were directed to the simulation-based practice, which took about 40 minutes to complete. During the simulation practice, participants were allowed to use all learning materials that were presented into the introduction module. After the simulation, participants completed the post-test to measure their knowledge and skills of the subject matter. The post-test took about 10 minutes to complete. Immediately after the post-test, participants completed both the engagement and demographic surveys. Figure 3 shows an overview of the procedure for this study.

Results

Using an analysis of variance (ANOVA), non-significant ($p > .05$) differences between the intervention groups were detected for gender and prior knowledge variables.

Testing internal consistency of the performance and motivation measures

The items of each of the scales used in the study were tested for Internal Reliability by using Cronbach's alpha coefficient. The 18 SDVVER items resulted in Cronbach's $\alpha = .82$ for pre-test, $\alpha = .68$ for post-test, and $\alpha = .80$ for test-retest. The 12-item IMMS measure consisted of 3 items for each subscale. The overall internal consistency was sufficient ($\alpha = .86$). The subscales shared moderate values of internal consistency ($\alpha = .72$ for attention, $\alpha = .73$ for relevance, $\alpha = .64$ for confidence, and $\alpha = .63$ for satisfaction).

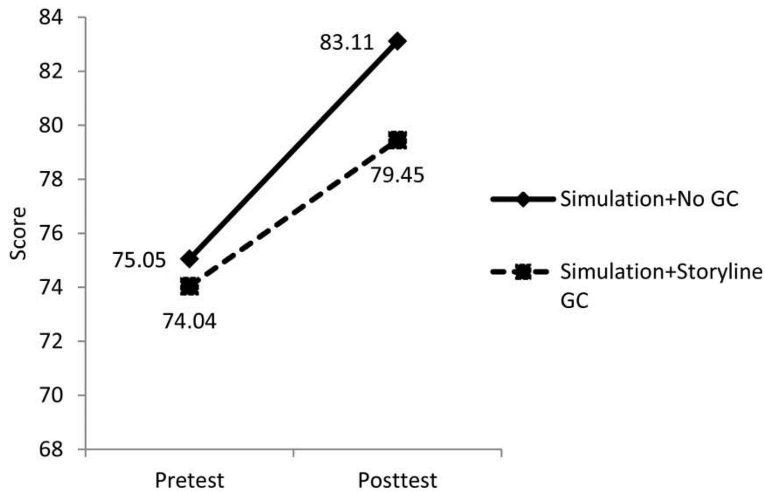


Figure 4. Means of learning effectiveness for pre- and post-tests by learning condition.

Learning effectiveness

To test the intervention's learning effectiveness, a repeated measures mixed analysis of variance (ANOVA) was performed. This analysis provides information on treatment effects (between groups), pre–post time effects (within groups), and their interaction (i.e. whether one intervention enhanced learning more than the other one). A significant time effect on learning effectiveness was obtained, Wilks' $\lambda = .77$, $F(1, 62) = 18.99$, $p < .001$, $\eta^2 = .23$. This effect indicates a significant improvement in learning effectiveness from the pre- to post-test across the two intervention groups ($M_{\text{Pre-test}} = 74.54$, SD = 16.89, $M_{\text{Post-test}} = 81.28$, SD = 13.05, $d = .45$). A non-significant intervention by time interaction was obtained, Wilks' $\lambda = .99$, $F(1, 62) = .73$, $p = .40$, $\eta^2 = .01$ (Figure 4), indicating that students in both interventions enhanced their statistical knowledge similarly over time. Furthermore, the intervention effect on learning effectiveness failed to reach significance, $F(1, 62) = .457$, $p = .50$, $\eta^2 = .02$, indicating insufficient evidence that learning effectiveness of Simulation+No GC and Simulation+Storyline GC interventions was different. However, there seems to be a moderate to small effect size between the two learning conditions ($M_{\text{No GC}} = 83.11$, SD = 12.91 vs. $M_{\text{GC}} = 79.45$, SD = 13.14, Cohen's $d = .28$) relative to the post-test score.

Learning efficiency

The intervention (i.e., learning conditions) effect on learning efficiency was tested by a one-way ANOVA. Non-significant differences between the two interventions on the time spent on simulation-based practice and learning efficiency were found, $F(1, 62) = 1.31$, $p < .26$, $\eta^2 = .02$, and $F(1, 62) = 2.74$, $p < .10$, $\eta^2 = .04$, respectively. The learning efficiency of the Simulation+No GC group was slightly higher than the Simulation+Storyline GC group, and there was a very small effect size ($M_{\text{No GC}} = 2.51$, SD = .78 vs. $M_{\text{Storyline GC}} = 2.18$, SD = .81, $d = .006$).

Learning engagement

Learning engagement (IMMS total score) was subjected to a one-way between subjects ANOVA. Significant differences between the two interventions emerged, $F(1, 61) = 4.76$,

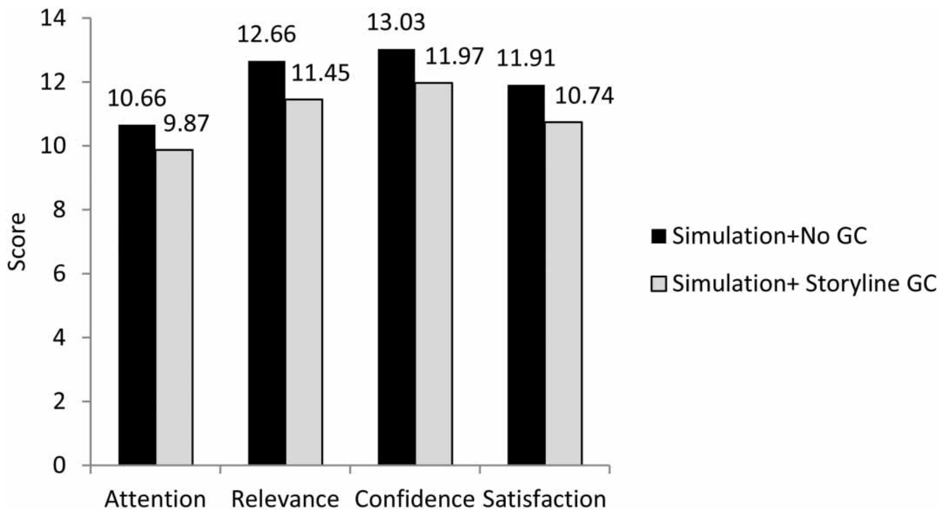


Figure 5. Means of ARCS by intervention condition.

$p < .03$, $\eta^2 = .07$. The overall learning engagement of students from the Simulation+No GC condition was significantly higher than the Simulation+Storyline GC group ($M_{\text{No GC}} =$ 48.25, SD = 8.25, $M_{\text{Storyline GC}} = 44.03$, SD = 7.02, $d = .55$).

The ARCS components of the IMMS survey were analyzed by using a multivariate ANOVA. The analysis revealed a non-significant effect of interventions on ARCS components, Wilks' $\lambda = .91$, $F(4, 58) = .40$, $p < .25$, $\eta^2 = .09$. However, separate univariate ANOVAs applied to these components revealed (1) significant intervention effects on the Relevance dimension, $F(1, 61) = 4.24$, $p < .04$, $\eta^2 = .06$ ($M_{\text{No GC}} = 12.66$, SD = 2.30 vs. $M_{\text{Storyline GC}} = 11.45$, SD = 2.35, $d = .52$), and the Satisfaction dimension, $F(1, 61) = 4.23$, $p < .04$, $\eta^2 = .06$ ($M_{\text{No GC}} = 11.91$, SD = 2.45 vs. $M_{\text{Storyline GC}} = 10.74$, SD = 2.02, $d = .52$), (2) non-significant interventions effect on Attention, $F(1, 61) = 1.19$, $p < .28$, $\eta^2 = .02$ ($M_{\text{No GC}} = 10.66$, SD = 2.95 vs. $M_{\text{Storyline GC}} = 9.87$, SD = 2.77, $d = .28$), and (3) approaching significance intervention effects on Confidence, $F(1, 61) = 3.62$, $\rho < .06$, $\eta^2 = .05$($M_{\text{No GC}} = 13.03$, SD = 2.06 vs. $M_{\text{Storyline GC}} = 11.97$, SD = 2.37, $d = .48$) (Figure 5).

In addition to filling out the IMMS survey, students reported how interesting they found the simulation context (i.e., the Career Coach storyline for the storyline-based group and analyzing job opportunities data for the non-storyline group) by answering a question on a 5-point Likert-type scale ranging from 1 (*not true at all*) to 5 (*very true*).

A one-way ANOVA revealed significant differences between the two interventions, $F(1, 52) = 9.60$, $p < .003$, $\eta^2 = .15$. Students from the Simulation+No GC found the context significantly more interesting than the Simulation+Storyline GC group ($M_{\text{No GC}} = 4.33$, SD = .92 vs. $M_{\text{Storyline GC}} = 3.54$, SD = .95, $d = .85$).

Discussion

One of the key findings of the study was a positive effect of simulation-based learning on statistics knowledge and skills acquisition for graduate students. Students' performance in both instructional methods improved from pre- to post-test ($d = .45$). However, an absence of a control condition, i.e., a group of learners who were not engaged in simulation-based learning, prevents us from generalizing this finding and concluding that the two

Table 2. Summary of the effects of storyline on learning effectiveness, efficiency, and engagement as compared to simulation without storyline.

Variables	Effect
Learning effectiveness	↓
Learning efficiency	↓
Learning engagement	↓*
Attention	↓
Relevance	↓*
Confidence	↓**
Satisfaction	↓*

Notes: ↑, positive effect; ↓, negative effect; =, no effect.
*Significant, $p < .05$.
**Approaching significance, $p = .06$.

interventions were the primary cause for this learning gain. However, based on previous research (de Croock & van Merrienboer, 2007), it is plausible that the two instructional methods had a positive effect on the learning gain.

Unfortunately, none of the research hypotheses postulating positive effects of a storyline on e^3-learning was supported by the data (Table 2). It is important to mention that learning effectiveness, efficiency, and engagement qualities of e^3-learning are interrelated. For example, learning efficiency depends on learning effectiveness, since students' performance is one of the parameters affecting learning efficiency. In addition, students' engagement may directly affect learning effectiveness since motivation is one of the factors influencing human performance (Barab, Gresalfi, & Ingram-Goble, 2010).

The assumption that a storyline coupled with a simulation would produce a larger effect on learning compared to a condition without a storyline was not supported by the data. These findings are in line with previous studies demonstrating non-significant performance differences between storyline and non-storyline instructional interventions (DeRouin-Jessen, 2008; Greenwood-Ericksen, 2008). The storyline added to a simulation may have distracted students from the instructional task, and posed a higher level of extraneous cognitive load that might be detrimental to learning outcomes. According to the Cognitive Load Theory (Sweller, 1994), information presentation method(s) can create unrelated/extraneous cognitive load. Due to limited cognitive working memory, if students need to process excessive information that is not directly related to the instructional task (i.e. following the storyline itself in addition to learning statistics), then the excessive information may create a deficit in cognitive working memory that is available for learning.

Examining statistics learning in a storyline context can be possibly related to mathematics story problems research that debates the pros and cons of placing mathematics in real-life contexts. Although the importance of relating instructional context to students' life experiences is widely acknowledged, it introduces many challenges to educators and students as well. Solving math word problems is a difficult task for many students since it requires both math skills and complex analytical strategies to create a situational model that represents story events, setting, and relationships (Nathan, Kintsch, & Young, 1992; Walkington et al., 2012). The study participants' job-related life experiences could be different from those presented in the storyline, which may partially explain why adding a storyline did not produce the expected results.

Adding a storyline may also inhibit learners' attention directed to the instructional task, and thus impinge on their learning. Attending to the storyline while trying to solve statistical problems might draw students' attention away from the learning process similarly to how talking and driving increase time spent on driving task and decreases visual processing speed (Puell & Barrio, 2008). Finally, it is plausible that significantly lower engagement levels reported by the Simulation+Storyline GC students could negatively affect students' performance because of reciprocal relationships between motivation and human performance (Kebritchi, Hirumi, & Bai, 2010).

Contrary to our assumption of a positive effect of a storyline on learning efficiency, defined in the current study as the ratio of students' post-test performance scores to the time spent on simulation-based activity, non-significant differences in learning efficiency were observed among the instructional methods. Since students in both instructional methods performed similarly and spent approximately the same amount of time on the instructional task (non-significant difference of 3.72 minutes on average), the ratio of students' post-test scores to the time spent on training did not differ between the two instructional methods.

The most surprising finding of this study was contrary to our expectation of a positive effect of a storyline on engagement, significantly higher engagement levels ($d = .55$) were observed among students learning with the Simulation+No GC than among students learning with the Simulation+Storyline GC. In spite of the significant differences in the engagement levels between the two conditions, the results suggest that both instructional methods were effective at engaging the participants on all four levels of motivation, i.e., attention, relevance, confidence, and satisfaction; reaching above 80% on the overall engagement scale for the Simulation+No GC instruction and above 73% for the Simulation+Storyline GC instruction.

There are several explanations of the storyline instructional method failing to increase students' engagement. First, the analysis of students' interest toward the simulation context (e.g. the "career coach" storyline for the storyline-based instruction, and "analyzing job opportunities data analysis" context for the non-storyline instruction) indicates that students learning with the simulation with no storyline found the context of analyzing job opportunities significantly more interesting than the students learning with the simulation with the storyline. Probably, the lower level of interest toward the storyline suggests a lower level of storyline relevance experienced by the learners, which was further supported by the lower relevance levels reported by the students learning with the Simulation+Storyline GC as compared with their Simulation+No GC counterparts. Furthermore, adding the storyline to the simulation also affected students' satisfaction. Students instructed by the Simulation+No GC method reported significantly higher satisfaction than the students instructed by the Simulation+Storyline GC method, which can be explained by the fact that the simulation did not consider students' preferences related to (1) the presence or absence of a storyline in the instructional activities, and (2) the storyline scenario context.

Designing relevant games and storylines, in particular, underscores one of the primary problems associated with game-based learning research. Researchers argue that one of the reasons for research in this area failing to provide hard evidence for the effectiveness of games as instructional tools is because most of the existing games are limited in responding to specific needs of individual students (Conati & Maclaren, 2009). Players experience different emotions while playing the same game. For example, a game that revolves around a storyline, or involves players exploring the game world, can be very appealing to students who are mainly interested in entertainment experiences. On the other hand, students whose primarily goal is to engage in learning through clearly defined activities may feel frustrated or anxious. The simulation used in this study was designed for graduate students. It is possible that the participants expected to go through traditional drill-and-practice materials that would help them acquire greater expertise in the subject matter rather than interact with storyline characters. One solution to this problem is to design games with adaptive instructional content and GCs that would provide students with the freedom to configure a game based on their own preferences.

Another possible issue that may explain the lackluster storyline effect on engagement is that the current instruction focused only on one GC. Usually, a game comprises 6–7 GCs where one GC can mediate another GC or a subset of GCs (Shute, Ventura, Kim, & Wang, in press). For example, the effect of a mild storyline may be improved by adding competition or collaboration mechanics to the game. This study focused only on one GC – storyline. Thus, if a competition characteristic was combined with a storyline, students might find the simulation-based materials more engaging. However, examining learning benefits of a game comprised of multiple GCs does not provide clear evidence for the reason that a particular game failed to support learning from a game-design standpoint. To design really effective games, we need to know how each GC and combination of GCs affect learning (Federation of American Scientists, 2006; Wilson et al., 2009).

In summary, developing instructional materials that evolve around a story is a nontrivial undertaking, requiring instructional designers to tie together an engaging storyline with instructional content. As Laaksolahti (2008) mentioned, if we want to create emotionally engaging storylines, we need to focus on the evaluation of experiences created by a story. This is a complicated effort, since most of the available instructional design tools are not suited for this task. The storyline developed for the present study was lacking in several ways from traditional storylines – such as building to climax and creating an emotional tension, which probably accounts for the limited desirable effects on engagement. Hence, the benefits and shortcomings of adding a storyline to a simulation should be carefully considered.

The digital gaming industry spends $10 billion per year on game development and production (Van Eck, 2006). The research evaluating the effects of GCs and storyline, in particular, on learning is an important component for improving game design, and subsequently instructional methods (Federation of American Scientists, 2006; Wilson et al., 2009). Future research should explore the effects of various GCs on learning outcomes. Specifically, researchers should examine games with adaptive GCs, i.e. game design. Games with adaptive game design may allow students to choose the desired level of challenge, completion, the reward system, interactivity, and the presence or absence of a storyline/fantasy.

Limitations

There were several limitations to the findings of this study. First, the instructional intervention developed for the study was relatively brief (i.e. 40 minutes) which could explain why adding a storyline did not result in significant time differences and consequently learning efficiency between the two conditions. In addition, only one storyline was developed, which limited students' ability to engage in a learning context of their choice.

Another limitation of the study was the small sample size. Using a larger sample size in future studies can provide more accurate results and increase the generalizability of the findings as well as allow for including a control group in the research design.

Conclusion

With regard to the development of storylines, even though there is no established recipe or formula on how to write successful stories, researchers and game developers need more specific guidelines on how to create storylines for a particular audience or learning context. Although the present study has not shown sufficient evidence for the positive effect of storylines in simulations, storylines do play an important role in human learning. According to the situated cognition perspective, all knowledge is situated in complex social systems (Greeno, 2006). Contextualized or storyline-enhanced problems offer many learning benefits, including improved confidence, creativity, and self-reliance (Walkington et al., 2012). Therefore, it is important to provide game developers and instructional designers with some guidelines on how to choose, develop, and embed a storyline into an educational game. Moreover, storylines or narratives may be used to enhance non-gaming instructional materials, e.g. books, simulations, case studies, and so on. As such, future research exploring effective strategies for presenting storyline-enhanced learning content is needed.

Notes on contributors

Elena Novak is an assistant professor in Instructional Design at Western Kentucky University. Her primary research interests focus on instructional games and simulations, and game-like learning environments. Some other research interests include digital storylines, emotionally intelligent pedagogical agents, and integration of technology into STEM learning. Elena Novak has been involved in several large-scale research projects that concerned technology integration in educational and military training settings.

Tristan E. Johnson is the Director of Online Education in the Graduate School of Engineering at Northeastern University. He has been involving in numerous R&D projects, including: leading the development of two online/hybrid engineering professional graduate degrees; the design and development of instructional and performance certifications for technical operations within Department of Defense and Department of Homeland Security; the design and development of online and blended learning solutions for US and International Universities; and the design and development of several technology support systems to support the use of the ISD methodology for developing complex cognitive training tasks as well as the use of a Navy-wide Human Performance Technology Support System.

Gershon Tenenbaum, PhD, Benjamin S. Bloom Professor of Educational Psychology, a graduate of Tel-Aviv University and the University of Chicago is a Professor of Sport and Exercise Psychology at the Florida State University, USA. He is a former director of the Ribstein Center for Research and Sport Medicine at the Wingate Institute in Israel, and coordinator of the Graduate Program in sport psychology at the University of Southern Queensland in Australia. From 1997 to 2001 he was the President of the International Society of Sport Psychology, and from1996 to 2008 he was the Editor of the *International Journal of Sport and Exercise Psychology*. He published extensively in psychology and sport psychology in areas of expertise and decision-making, psychometrics, and coping with physical effort experiences. Gershon received several distinguished awards for his academic and scientific achievements, and is a member and fellow of several scientific and professional forums and societies.

Valerie J. Shute is the Mack & Effie Campbell Tyner Endowed Professor in Education in the Department of Educational Psychology and Learning Systems at Florida State University. Her current research involves using immersive games with stealth assessment to support learning of cognitive and non-cognitive knowledge, skills, and dispositions. Her research has resulted in numerous grants, journal articles, books, chapters in edited books, and a patent.

References

Ang, C. S., Zaphiris, P., & Mahmood, S. (2007). A model of cognitive loads in massively multiplayer online role playing games. *Interacting with Computers*, *19*(2), 167–179.

Azevedo, R. (2005). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist*, *40*, 193–197.

Barab, S. A., Gresalfi, M., & Ingram-Goble, A. (2010). Transformational play: Using games to position person, content, and context. *Educational Researcher*, *39*, 525–536.

Barab, S. A., Zuiker, S., Warren, S., Hickey, D., Ingram-Goble, A., Kwon, E.-J., … Herring, S. C. (2007). Situationally embodied curriculum: Relating formalisms and contexts. *Science Education*, *91*(5), 750–782. doi:10.1002/sce.20217

Chang, K.-E., Chen, Y.-L., Lin, H.-Y., & Sung, Y.-T. (2008). Effects of learning support in simulation-based physics learning. *Computers and Education, 51*, 1486–1498.

Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, *19*, 267–303.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, *88*, 715–730.

de Croock, M. B. M., & van Merrienboer, J. J. G. (2007). Paradoxical effects of information presentation formats and contextual interference on transfer of a complex cognitive skill. *Computers in Human Behavior*, *23*, 1740–1761.

DeRouin-Jessen, R. E. (2008). *Game on: The impact of game features in computer-based training* (Unpublished PhD dissertation). University of Central Florida, Orlando, FL.

Dickey, M. D. (2006). Game design narrative for learning: Appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. *Educational Technology Research and Development*, *54*, 245–263.

Federation of America Scientists. (2006). Summit on educational games: Harnessing the power of video games for learning. Washington, DC: Author.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Gee, J. P. (2008). Learning and Games. In K. Salen, D. The John, & T. Catherine (Eds.), *The ecology of games: Connecting youth, games, and learning* (pp. 21–40). MacArthur foundation series on digital media and learning. Cambridge, MA: The MIT Press.

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, *14*(4), 311–327.

Greeno, J. (2006). Learning in activity. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 79–96). St. Louis, MO: Cambridge University Press.

Greenwood-Ericksen, A. (2008). *Learning African-American history in a synthetic learning environment* (Unpublished dissertation). University of Central Florida. Retrieved from http://search.proquest.com/docview/89208380?accountid=4840

Habgood, M. P. J., Ainsworth, S. E., & Benford, S. (2005). Endogenous fantasy and learning in digital games. *Simulation Gaming*, *36*, 483–498.

Huang, W., Huang, W., Diefes-Dux, H., & Imbrie, P. K. (2006). A preliminary validation of attention, relevance, confidence and satisfaction model-based instructional material motivational survey in a computer-based tutorial setting. *British Journal of Educational Technology*, *37*(2), 243–259.

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education, 55*, 427–443.

Keller, J. M. (1993). *Motivation by design*. Manuscript. Tallahassee, FL: Florida State University.

Keller, J. M. (2008). First principles of motivation to learn and e^3-learning. *Distance Education*, *29*, 175–185.

Laaksolahti, J. (2008). *Plot, spectacle and experience: Contributions to the design and evaluation of interactive storytelling* (PhD dissertation). Stockholm University.

Lee, C.-Y., & Chen, M.-P. (2009). A computer game as a context for non-routine mathematical problem solving: The effects of type of question prompt and level of prior knowledge. *Computers & Education*, *52*, 530–542.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science*, *4*, 333–369.

Marchiori, E. J., Torrente, J., Blanco, Á. d., Moreno-Ger, P., Sancho, P., & Fernández-Manjón, B. (2012). A narrative metaphor to facilitate educational game authoring. *Computers & Education*, *58*, 590–599.

McKee, R. (2005). *Story – substance, structure, style and the principles of screenwriting*. York: Methuen Publishing.

Mori, Y., Yamamoto, Y., & Yadohisa, H. (2003). *Data-oriented learning system of statistics based on analysis scenario/story (DoLStat)*. Paper presented at the Bulletin of the International Statistical Institute, 54th session proceedings. Berlin, Germany.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction*, *9*(4), 329–389.

Parker, L. E., & Lepper, M. R. (1992). Effects of fantasy contexts on children's learning and motivation: Making learning more fun. *Journal of Personality and Social Psychology*, *62*, 625–633.

Puell, M. C., & Barrio, A. (2008). Effect of driver distraction and low alcohol concentrations on useful field of view and frequency-doubling technology perimetry. *Acta Ophthalmologica*, *86*, 634–641.

Rouse, R. (2005). *Game Design: Theory and Practice*. Plano, TX: Jones & Bartlett.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*, 153–189.

Shute, V. J., & Glaser, R. (1990). Large-scale evaluation of an intelligent tutoring system: Smithtown. *Interactive Learning Environments, 1*, 51–76.

Shute, V. J., & Ke, F. (2012). Games, learning, and assessment. In D. Ifenthaler, D. Eseryel, & Ge, X. (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 43–58). New York, NY: Springer.

Shute, V. J., Ventura, M., Kim, Y. J., & Wang, L. (in press). Video games and learning. In W. G. Tierney, Z. Corwin, T. Fullerton & G. Ragusa (Eds.), *Postsecondary Play: The Role of Games and Social Media in Higher Education*. Baltimore, MD: John Hopkins University Press.

Song, S., & Keller, J. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. *Educational Technology Research and Development*, *49*, 5–22.

Spires, H. A., Turner, K. A., Rowe, J., Mott, B., & Lester, J. (2010). *Game-based literacies and learning: Towards a transactional theoretical perspective*. Paper presented at the American Educational Research Association (AERA), Denver, CO.

Squire, K. (2005). *Game-based learning: Present and future state of the field*. Houston, Texas: Masie Center e-Learning Consortium.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295–312.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review*, *41*, 16–30.

Walkington, C., Sherman, M., & Petrosino, A. (2012). "Playing the game" of story problems: Coordinating situation-based reasoning with algebraic representation. *The Journal of Mathematical Behavior*, *31*(2), 174–195. doi:10.1016/j.jmathb.2011.12.009

Watson, W. R., Mong, C. J., & Harris, C. A. (2011). A case study of the in-class use of a video game for teaching high school history. *Computers & Education*, *56*, 466–474.

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E. C., Burke, S., Estock, J. L., et al. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming*, *40*(2), 217–266.