Toward Understanding the Impact of Visual Themes and Embellishment on Performance, Engagement, and Self-Efficacy in Educational Games

Toward Understanding the Impact of Visual Themes and Embellishment on Performance, Engagement, and Self-Efficacy in Educational Games

Dominic Kao and D. Fox Harrell
Massachusetts Institute of Technology

Kao, D., & Harrell, D. F. (2017). Toward Understanding the Impact of Visual Themes and Embellishment on Performance, Engagement, and Self-Efficacy in Educational Games. Paper presented at annual American Educational Research Association meeting, San Antonio, TX, USA.

Abstract

The results of over twenty-five years of research seem clear: the addition of seductive visual details in video games hinders performance of learners (Garner, Gillingham, & White, 1989; Thalheimer, 2004; Rey, 2012). Yet, countless other research results propose the opposite: that visual embellishments and well-designed ambiguity instead *improve* learners' performance, engagement, and self-efficacy (Tierney, Corwin, Fullerton, & Ragusa, 2014; Wilson et. al 2009; Scott & Ghinea, 2013). To shed light on this apparent contradiction, we devised a particular experiment using *game skins* to implement variations in visual themes of a computer game. Game skins are coherent, interchangeable sets of graphical assets that all implement the same underlying game structure while varying the visual appearance (for instance, see Figure 3). In particular, we implemented the following four game skins labeled and described as follows: 1) *Generic* theme with no embellishments (simple flat color background), 2) *Fantasy* game theme (forest, snow, and desert adventure backgrounds), 3) *STEM-oriented* theme (computer circuitry background), and 4) *Choice* (the user picks one of the previous three

options). Our goal is determining if there are differences in performance, engagement, and self-efficacy between conditions. The upshot is that the generic condition participants had highest performance (levels) and had highest programming self-efficacy—followed by choice, fantasy game setting, circuitry. However, ordering of conditions for engagement was precisely opposite the trend for performance. We conclude by discussing the trade-offs between the two diametrically opposed approaches to game themes and embellishment: instrumental game skins vs. thematic and deliberately embellished game skins.

Motivation

One of the largest paradigm shifts in the last thirty years has been movement away from the learning as an acquisition metaphor (Sfard, 1998) and instead toward a concept of learning as fundamentally contextually situated (Lave, 1991; Lemke, 1997; Walkerdine, 1997; Wenger, 1998; Barab & Duffy, 2000; Greeno, 1998; Lave & Wenger, 1991; Resnick, 1987; Young, 1993). One resultant argument is that people develop deep expertise—islands of expertise—that then lead to the formation of overarching themes, abstract enough that they engender further learning both within and outside of the original topic of interest (Crowley & Jacobs, 2002; Shaffer, 2006). Given the vast proliferation of educational games, adaptive learning systems, and Massive Open Online Courses (MOOCs) in recent years (Yuan & Powell, 2013), it is increasingly important to understand the significance how educational content is situated within computer-based learning environments (Wilson et. al, 2009; Garris, Ahlers, & Driskell, 2002; Malone & Lepper, 1987; Rieber, 1996; Driskell & Dwyer,

1984), e.g., ranging from STEM-oriented to fantasy settings in educational games. For decades researchers have found that embellishing instruction with fantasy content, improves instructional efficacy, e.g., as in (Cordova & Lepper, 1996; Scott & Ghinea, 2013; Asgari & Kaufman, 2004; Daanen & Grant, 2007; Resnick et. al, 2009). Games are touted to move beyond the "content fetish" (Gee, 2008) so prevalent in society and to immerse players in an experience where there is *intentional* inefficiency in conveyed content. That is, instead of trying to rush toward "instrumentalized" games (Zimmerman, 2011), it is specifically the embellished ambiguities that create opportunities to explore (Fullerton, 2014).

Yet, in making this argument we need to account for the fact that this is the opposite of what some researchers in the learning sciences would postulate. The opposing viewpoint holds that that such embellishments would constitute *seductive details* that impede educational efficacy (Garner, Gillingham, & White, 1989; Thalheimer, 2004; Rey, 2012; Park, Moreno, Seufert, & Brünken, 2011; Dewey, 1913; Sanchez & Wiley, 2006; Harp & Mayer, 1998; Rummer, Schweppe, Fürstenberg, Seufert, & Brünken, 2010; Lehman, Schraw, McCrudden, & Hartley, 2007). The *coherence principle* of multimedia learning is a culmination of this line of work. It advises removing any illustration not of fundamental importance to the instructional goal (Mayer, 2005; Clark & Mayer, 2011).

Here, our goal is to explore and investigate these opposing viewpoints. We consider how three different game skins affect participants' performance, engagement, and self-efficacy. We find that the more embellished and more ambiguous, game skins thwart performance, but *improve* engagement. Our results

suggest that simpler game skins improve performance, but *reduce* engagement. Such a trade-off is particularly important in educational games, in which both performance *and* engagement are highly desirable to the end goal (Harteveld, 2015; Blumenfeld, Kempler, & Krajcik, 2005). We conclude with a reflective discussion on how educators and developers might navigate this dual goal.

The Game

The experiment takes place in a STEM learning game called *Mazzy* (Kao & Harrell, 2015e).[1] *Mazzy* is a game in which players complete maze-like challenges by creating short computer programs. In total, there are 12 levels in this version of *Mazzy*. Levels 1-5 require only basic commands. Levels 6-9 require using loops. Levels 10-12 require using all preceding commands in addition to conditionals (see Figures 1 and 2). *Mazzy* has been used previously as an experimental testbed for evaluating the impacts of avatar type on performance and engagement in an educational game (Kao & Harrell, 2015a-d; Kao & Harrell, 2016a-d).

Methods

Our experiment compared the impacts of four game skin conditions: (a) Generic Theme, (b) Fantasy Theme, (c) STEM Theme (circuit board), and (d) User Choice. The goal was to see if participants using different game skins vary in performance, engagement, and self-efficacy. We suspected that (1) the generic skin would have

[1] Gameplay video: http://youtu.be/n2rR1CtVal8

the highest performance, but that (2) the embellished skins would have the highest engagement.

Game Skin Conditions

The four game skin conditions we tested were:

a. Generic Theme
b. Fantasy Theme
c. STEM Theme
d. User Choice

The generic theme was specifically made to have no embellishments, just flat color. The fantasy theme and STEM (circuitry) theme were heavily embellished in their respective themes (see Figures 3, 4 and 5). A choice condition was included to test if users given choice of game skin have improved performance (Ryan & Deci, 2000; Martin et. al, 2003; Jolivette et. al 2001; Hardre & Reeve, 2003; Flowerday & Schraw, 2000; Eshel & Kohavi, 2003; Deci & Ryan, 1985; Becker, 1997; Assor et. al, 2002; Kernan et. al, 1991; Patall, et. al, 2008; Cordova & Lepper, 1996). This lattermost condition begins with players selecting a game skin—choices always appear in a random order—afterwards all aspects of the game are exactly identical. See Figure 6. The player avatar is a blue triangle (Munsell color *7.5PB 5/18*)—the avatar color was chosen to minimize interaction effects with game skins. This was later checked post-game, e.g., virtually all players irrespective their condition (given a range of *1: Strongly Disagree* to *5: Strongly Agree*) strongly disagreed that the avatar clashed with the background (M=1.46, SD=0.89).

Quantitative and Qualitative Measures

Performance was measured as a function of levels completed, number of attempts, and number of hints. Engagement was measured using the Player Experience of Needs Satisfaction (PENS) scale (Ryan, Rigby, & Przybylski, 2006) and the Game Experience Questionnaire (GEQ) (IJsselsteijn, Kort, Poels, Jurgelionis, & Bellotti, 2007). Self-Efficacy was measured using the Computer Programming Self-Efficacy Scale (CPSES) (Ramalingam, & Wiedenbeck, 1998). Our instrument was a selected portion of the original CPSES scale. Principal components analysis (PCA) was performed to assess construct validity, with high validity metrics; reliability using Cronbach's alpha was also high, 94.4 percent. See Table 1.

Participants

1172 participants were recruited through Mechanical Turk (demographics Table 2). Participants were reimbursed $1.50 to participate in this experiment.

Design

A between-subjects design was used: game skin condition was the between-subject factor. Participants were randomly assigned to a condition.

Protocol

Prior to starting the game, players were informed that they could exit the game at *any time* via a red button in the corner of the screen. When participants were done playing (either by exiting early, or by finishing all 12 levels), participants returned

to the experiment instructions, which then prompted them with PENS, GEQ, and CPSES, then a demographics survey.

Analysis

Data was analyzed in SPSS using MANOVA. The dependent variables are levels completed, number of attempts, number of hints, and the PENS, GEQ, and CPSES; the independent variable is game skin condition. All the dependent variables are continuous variables. The independent variable game skin condition (i.e., 0=generic, 1=fantasy, 2=circuitry, 3=choice) was a quadchotomous variable. A MANOVA was run for performance and for each questionnaire. Before running MANOVAs, all the variables included in the analyses were checked. There were univariate outliers and also multivariate outliers, but no outlier was statistically significant so they were retained. One participant was removed for investing minimal effort (0 attempts, 0 levels completed). Prior to running our MANOVAs, we checked both assumption of homogeneity of variance and homogeneity of covariance by the test of Levene's Test of Equality for Error Variances and Box's Test of Equality of Covariance Matrices. Levene's test was met by the data ($p>.05$), but Box's test ($p<.05$) was found untenable. To address this violation, Pillai's Trace was used instead of Wilk's Lambda.

Results

Both embellishment and ambiguity appear to improve engagement but decrease performance. Performance was ordered: generic > choice > fantasy > STEM. Self-efficacy was ordered the same. Engagement was ordered: STEM > fantasy > choice > generic. This was consistent across several measures. The following lists describe these results in terms of performance, self-efficacy, and engagement in fuller detail.

Performance

- Average playtime 21.2 minutes—no notable differences across conditions.
- Overall MANOVA was significant, $p<0.001$ (Table 3).
- Univariate tests found all measures to be significant, $p<0.05$ (descriptives Table 4, posthocs Table 5).
- Across all performance measures, performance was consistently ordered: generic > choice > fantasy > STEM (see Figure 7).
- Moreover, this effect was found to be true throughout the entire game.

Self-Efficacy

- Overall MANOVA was significant, $p<0.05$ (Table 6).
- Univariate tests found eight (of twelve) CPSES questions to be significant, $p<0.05$ (descriptives Table 7, posthocs Table 8).
- On average, similar ordering to performance: generic > choice > fantasy > STEM (see Figure 8).

Engagement (GEQ)

- Overall MANOVA was significant, $p<0.001$ (Table 9).
- Univariate tests found eighteen GEQ questions to be significant, $p<0.05$ (descriptives Table 10, posthocs Table 11).
- On average, engagement was ordered: STEM > fantasy > choice > generic (see Figure 9).

Engagement (PENS)

- Overall MANOVA was significant, $p<0.001$ (Table 12).
- Univariate tests found six PENS questions to be significant, $p<0.05$ (descriptives Table 13, posthocs Table 14).
- Consistently, across all questions on autonomy, relatedness, and presence, conditions were ordered: STEM > fantasy > choice > generic (see Figure 10).

Choice

- Choice had *no* notable influence on performance, self-efficacy, engagement.
- True even when accounting for the skewed distribution of choices—generic (25%), fantasy (52%), STEM (23%) (descriptives Table 15).
- One potential explanation is that the choice presented was not very meaningful to participants (Flowerday & Schraw, 2000; Katz & Assor, 2007; Rose & Meyer, 2002; Evans & Boucher, 2015).

Discussion

Here, we discuss the importance of our findings, why they may have arose, and reflect on how developers and educators might navigate the trade-offs involved in two diametrically opposed approaches to game themes and embellishment.

We first summarize our findings:

- Generic skin condition participants had highest performance
- Generic skin condition participants had highest self-efficacy
- STEM/Fantasy condition participants had highest engagement

Why is this important? Games are clearly becoming ubiquitous—in 2015, the Entertainment Software Association (ESA) estimates that 155 million Americans play video games, 4/5 U.S. households own a device used to play video games, and 42% of Americans play video games regularly (3 hours or more per week) (ESA, 2015). Moreover, educators are increasingly trying to harness the potential of games for education; embedding content in fantasy settings is quickly becoming pervasive (Cordova & Lepper, 1996; Scott & Ghinea, 2013; Asgari & Kaufman, 2004; Daanen & Grant, 2007; Resnick et. al, 2009; Harrell & Veeragoudar Harrell, 2009; Gee, 2008). This approach has also been commercialized, e.g., Classcraft (2016), CodeCombat (2016), etc. However, developers' knowledge of how such embellishments may affect users in game-like environments is lacking. In the study reported on here, we found that embellishments may have significant effects on user performance, engagement, and programming self-efficacy. The implications are

important, e.g., self-efficacy is a strong predictor of women's career choices, especially in regards to STEM fields (Betz & Hackett, 1986; Pajares, 1996; Brauner, Leonhardt, Ziefle, & Schroeder, 2010). Moreover, performance *and* engagement are measures strongly correlated with learning and motivation (Harteveld, 2015; Blumenfeld et. al 2005). Thus, levels of embellishment appear to significantly influence users on a wide variety of crucial constructs.

Why did this happen? We posit that one cause is *seductive details*, which interfere with problem solving abilities in high cognitive load environments (Park et. al 2011; Park, Flowerday, & Brünken, 2015). This happens because of three things (Harp & Mayer, 1998): *distraction* (taking attention away from the relevant and moving it towards the irrelevant) (Sanchez & Wiley, 2006), *disruption* (making it harder to create correct mental schemas) (Lehman et. al, 2007), and *diversion* (priming prior knowledge that is unhelpful) (Rowland, Skinner, Davis-Richards, Saudargas, & Robinson, 2008; Harp & Mayer, 1998). This is well-known in instructional media, where embellishment is known to distract and also create ambiguity (e.g., line sketches vs. 3D graphics) (Butcher, 2006; Mayer, Hegarty, Mayer, & Campbell, 2005; Scheiter, Gerjets, Huk, Imhof, & Kammerer, 2009; Mayer, Heiser, & Lonn, 2001). Yet some researchers argue that embellishment has motivational affordances (Park et. al, 2011; Goetz & Sadoski, 1995; Ozdemir & Doolittle, 2015; Fullerton, 2014). Our results provide validity to both arguments—in our study comparing game skins, our results suggest that embellished themes may reduce performance all the while *improving* participant engagement.

What should developers do now? The implications are powerful. That the mere graphical skin of a game can impact users in a variety of important ways means that we can no longer simply assume that embellishing in fantasy is necessarily positive, e.g., (Garris & Ahlers, 2001; Parker & Lepper, 1992; Rieber, 1996; Westrom & Shaban, 1992; Virvou, Katsionis, & Manos, 2005; Wilson et. al, 2009), nor the inverse. Instead, we advocate to view embellishment holistically. In considering literature from different research fields, multiple, seemingly dichotomous perspectives are reconcilable under the tenet that *no global maximum exists*. Embellishment may affect performance adversely, all the while affecting engagement beneficially.

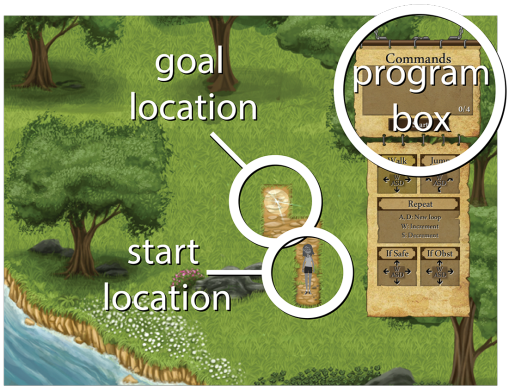
Our results also suggest another path forward. Developers must invest in compelling and coherent design. We can imagine a type of theme or skin that is elegant, imaginative, and domain-coherent that is a type of *best of both worlds* theme that would lead to high levels of both performance *and* engagement—themes that avoid unnecessary complexity and embellishment while maintaining elegant thematic coherence. In the future, we hope to further untangle the complicated constructs involved in assessing visual themes. Ultimately, such studies may be valuable for educational designers when it comes to creating diverse types of computer-based environments for learning.

Acknowledgements

We thank the anonymous reviewers for their valuable feedback. This work is supported by NSF STEM+C Grant #1542970 and an NSERC fellowship.

Figure 1 & 2

Mazzy Screenshots



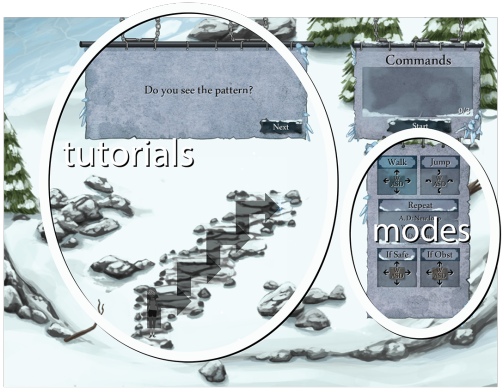


Figure 3

Levels 1-4

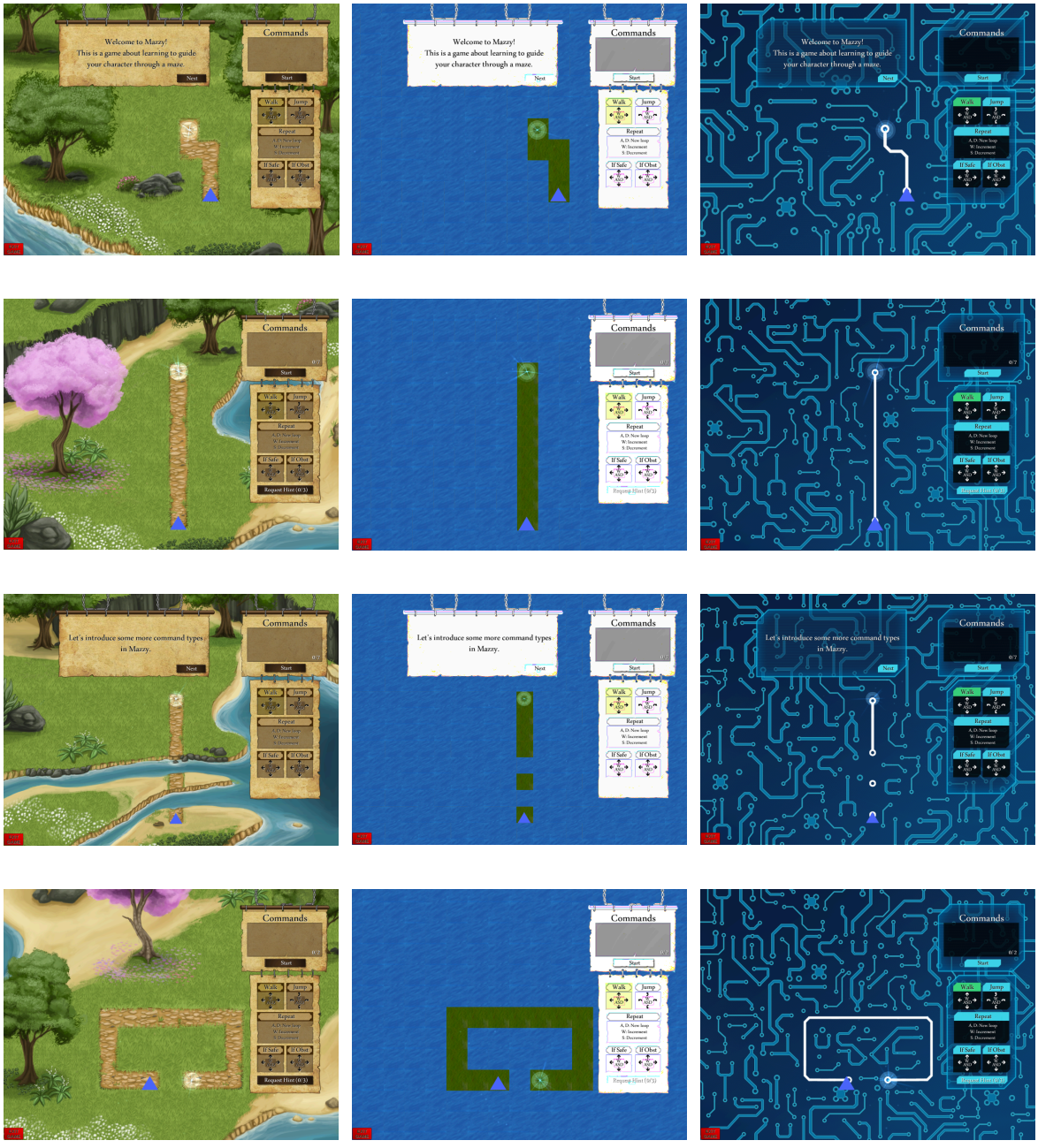


Figure 4

Levels 5-8

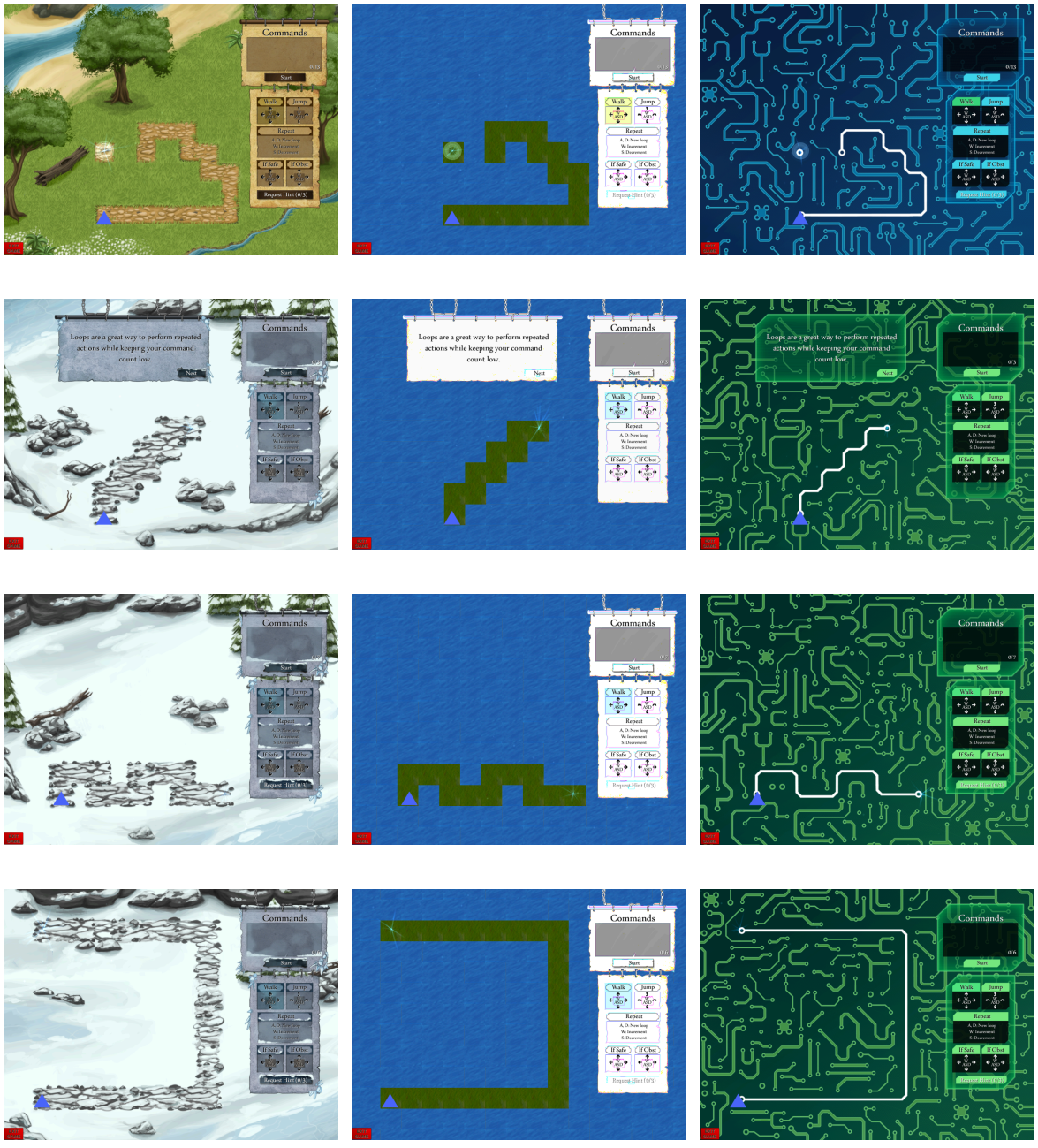


Figure 5

Levels 9-12

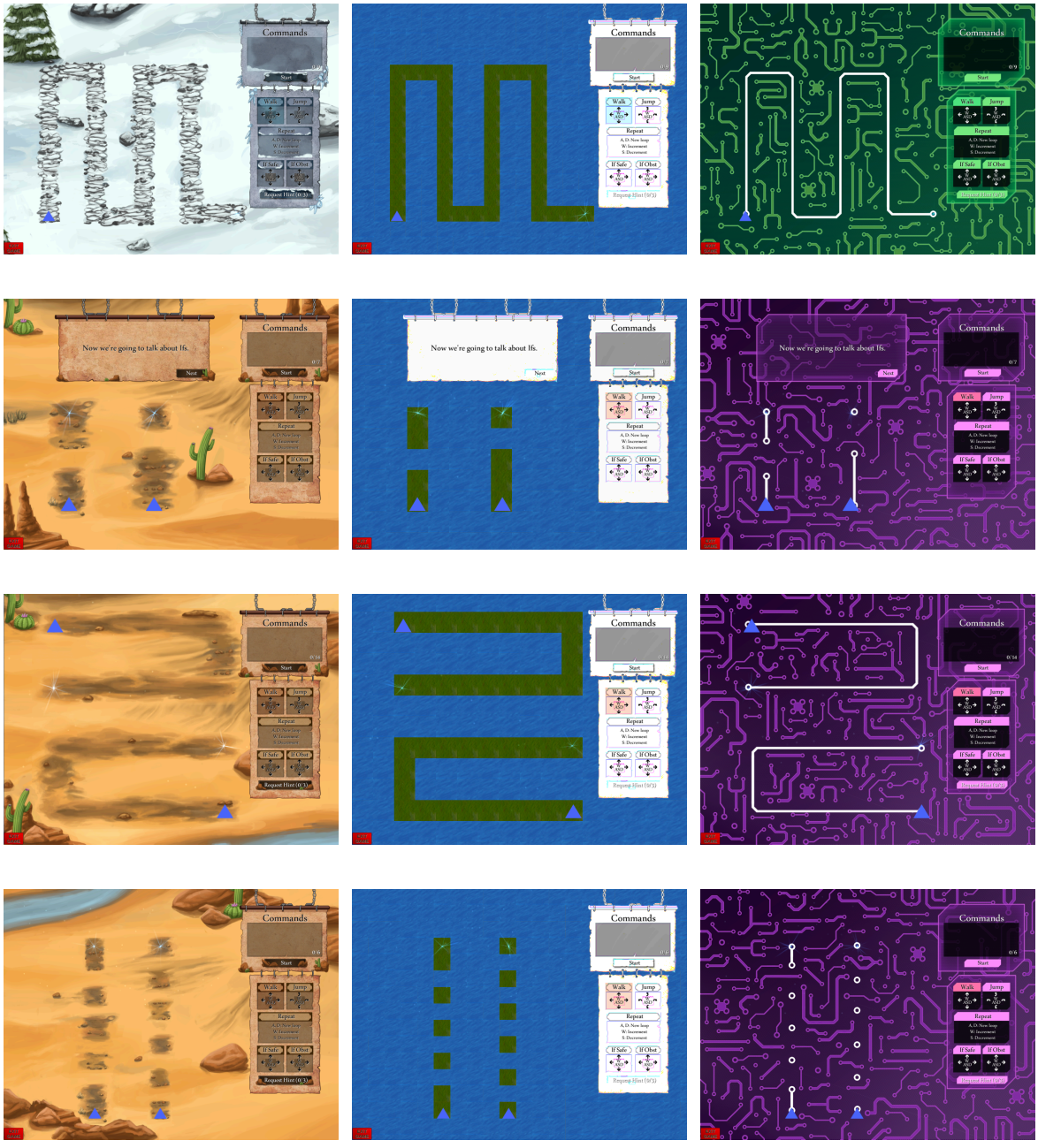


Figure 6

Choice Condition

Figure 7

Performance—Graphs

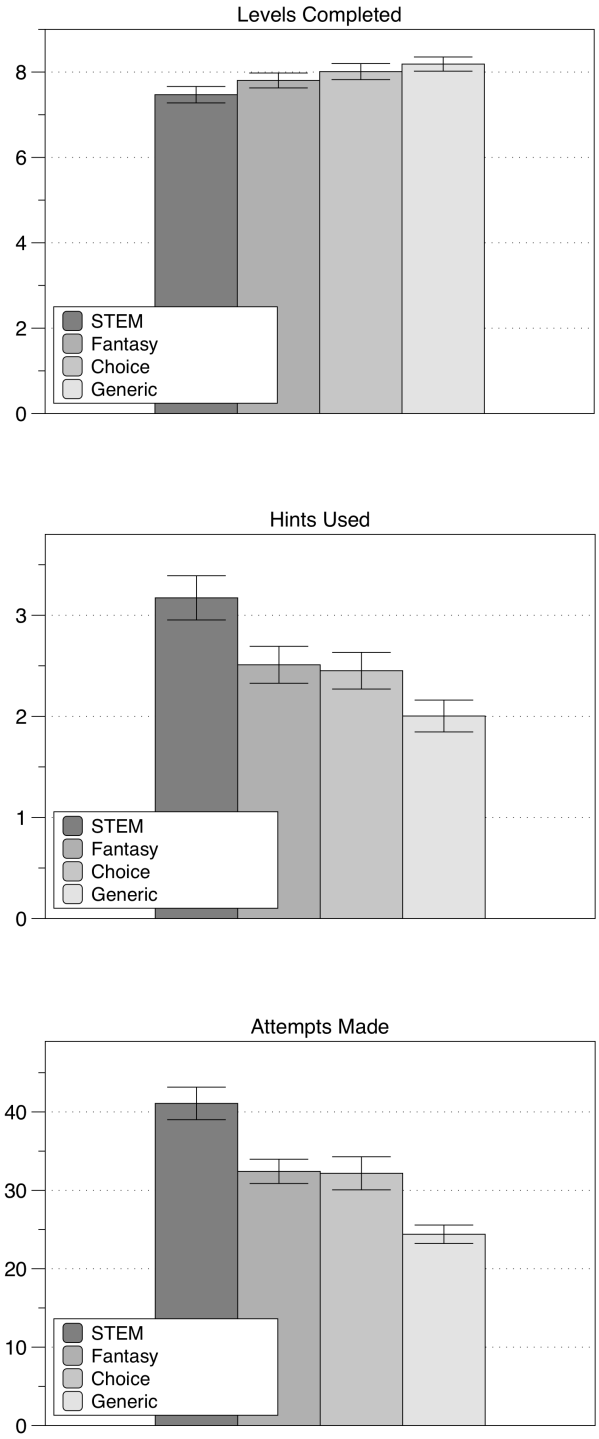


Figure 8

Self-Efficacy—Graph

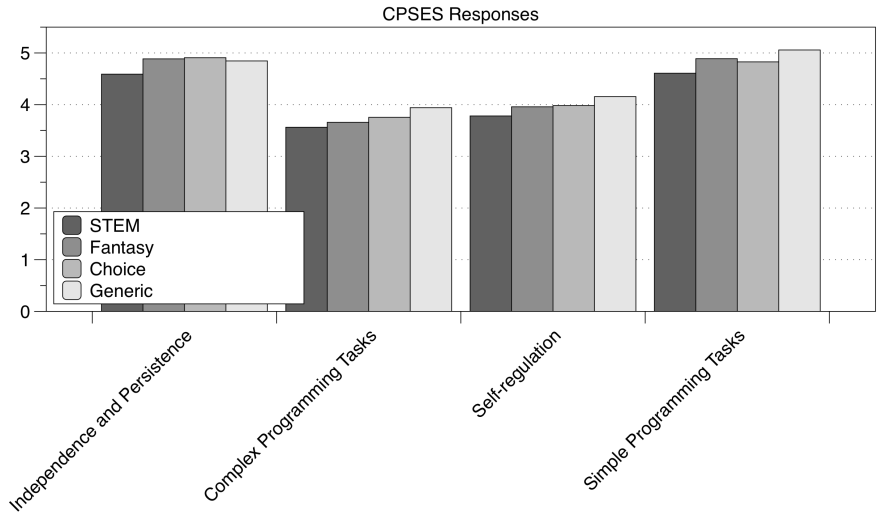


Figure 9

GEQ—Graph

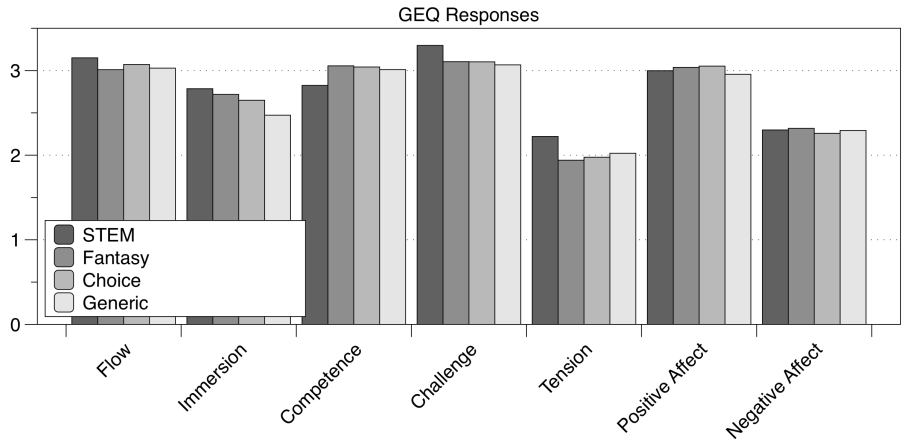


Figure 10

PENS—Graph

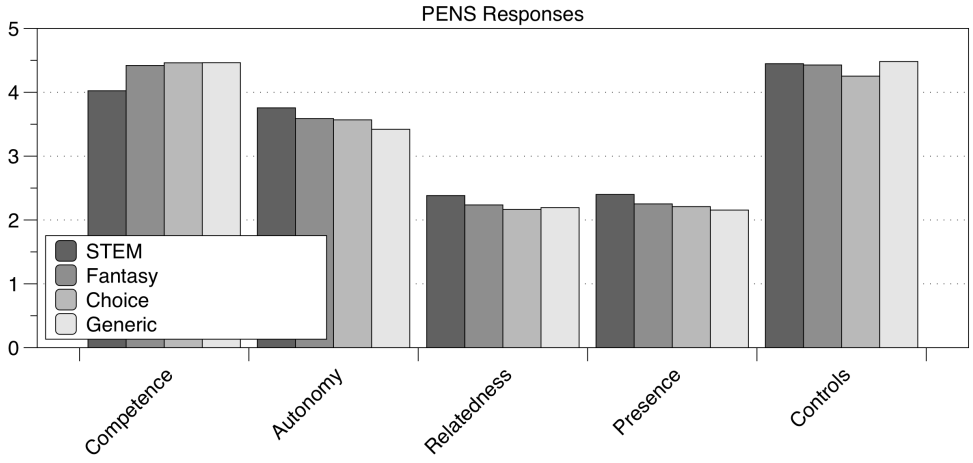


Table 1

CPSES

Factor 1: Independence and persistence (alpha = 0.84)

1. Complete a program if I had no help at all.
2. Complete a program once the tutorial helped me get started.
3. Complete a program if someone showed me how to solve the problem first.

Factor 2: Complex programming tasks (alpha = 0.85)

1. Write a program for an extremely difficult problem.
2. Organize my program in a clean way.
3. Mentally trace through the execution of a long, complex, program given to me.

Factor 3: Self-regulation (alpha = 0.85)

1. Come up with a suitable strategy for a given problem in a short time.
2. Manage my time efficiently if I had a pressing deadline on a problem.
3. Find a way to concentrate on my program, even when there were many distractions around me.

Factor 4: Simple programming tasks (alpha = 0.86)

1. Write logically correct blocks of code.
2. Write a program for a simple problem.
3. Write a program for a moderately difficult problem.

Table 2

Demographics

Characteristic	Category	n	%
Gender	Female	480	41.0
	Male	692	59.0
Age	18-20	73	6.2
	21-30	649	55.4
	31-40	308	26.3
	41-50	99	8.5
	>50	43	3.7
Ethnicity	White	944	80.6
	Black or African American	66	5.6
	Asian Indian	40	3.4
	Chinese	18	1.5
	Korean	8	0.7
	American Indian	9	0.8
	Vietnamese	7	0.6
	Japanese	3	0.3
	Filipino	5	0.4
	Other	72	6.1

1172 participants were recruited through Mechanical Turk. The data set consisted of 692 male, and 480 female participants. Participants self-identified their races/ethnicities as white (944), black or African American (66), Asian Indian (40), Chinese (18), Korean (8), American Indian (9), Vietnamese (7), Japanese (3), Filipino (5), and other (72). Participants were between the ages of 18 and 71 (M=30.1, SD=8.8), and were all from the United States. Participants were reimbursed $1.50 to participate in this experiment.

Table 3

Performance—MANOVA Multivariate F-tests

Effect		Value	*F*	Hypothesis *df*	Error *df*	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.871	2615.055[a]	3.000	1165.000	.000	.871
	Wilks' Lambda	.129	2615.055[a]	3.000	1165.000	.000	.871
	Hotelling's Trace	6.734	2615.055[a]	3.000	1165.000	.000	.871
	Roy's Largest Root	6.734	2615.055[a]	3.000	1165.000	.000	.871
NumericCondition	Pillai's Trace	.092	12.344	9.000	3501.000	.000	.031
	Wilks' Lambda	.908	12.777	9.000	2835.454	.000	.032
	Hotelling's Trace	.102	13.128	9.000	3491.000	.000	.033
	Roy's Largest Root	.101	39.277[b]	3.000	1167.000	.000	.092

a. Exact statistic
b. The statistic is an upper bound on F that yields a lower bound on the significance level.
c. Design: Intercept + NumericCondition

Table 4

Performance—Descriptive

				95% Confidence Interval	
Dependent Variable	Condition	Mean	Std. Error	Lower Bound	Upper Bound
Levels Completed	Generic	8.187	.178	7.837	8.536
	Choice	8.010	.182	7.654	8.367
	Fantasy	7.803	.177	7.456	8.150
	Circuit	7.470	.185	7.107	7.832
Hints Requested	Generic	2.003	.183	1.644	2.362
	Choice	2.451	.187	2.085	2.818
	Fantasy	2.510	.182	2.153	2.867
	Circuit	3.172	.190	2.800	3.544
Attempts	Generic	24.397	1.730	21.002	27.791
	Choice	32.174	1.766	28.709	35.638
	Fantasy	32.421	1.719	29.049	35.793
	Circuit	41.086	1.794	37.566	44.606

Table 5

Performance—Posthocs

Dependent Variable	Conditions	p-value
Levels Completed	Generic > Circuit	$p < .005$
Levels Completed	Choice > Circuit	$p < .05$
Hints Requested	Generic < Fantasy	$p < .05$
Hints Requested	Generic < Circuit	$p < .001$
Hints Requested	Choice < Circuit	$p < .01$
Hints Requested	Fantasy < Circuit	$p < .05$
Attempts	Generic < Choice	$p < .001$
Attempts	Generic < Fantasy	$p < .001$
Attempts	Generic < Circuit	$p < .001$
Attempts	Choice < Circuit	$p < .001$
Attempts	Fantasy < Circuit	$p < .001$
Attempts	Fantasy < Circuit	$p < .001$

Post hoc comparisons (LSD) revealed that participants in the generic condition completed more levels than participants in the circuit condition, $p<.005$. Participants in the choice condition also completed more levels than participants in the circuit condition, $p<.05$. Participants in the generic condition used less hints than participants in either the fantasy, $p<.05$, or circuit, $p<.001$, conditions. Participants in the choice condition used less hints than participants in the circuit condition, $p<.01$. Participants in the fantasy condition used less hints than participants in the circuit condition, $p<.05$. Participants in the generic condition used less attempts than participants in the choice, fantasy, or circuit conditions, $p<.001$. Participants in the choice condition used less attempts than participants in the circuit condition, $p<.001$. Participants in the fantasy condition used less attempts than participants in the circuit condition, $p<.001$.

Table 6

Self-Efficacy—MANOVA Multivariate F-tests

Effect		Value	*F*	Hypothesis *df*	Error *df*	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.934	1358.838[a]	12.000	1156.000	.000	.934
	Wilks' Lambda	.066	1358.838[a]	12.000	1156.000	.000	.934
	Hotelling's Trace	14.106	1358.838[a]	12.000	1156.000	.000	.934
	Roy's Largest Root	14.106	1358.838[a]	12.000	1156.000	.000	.934
NumericCondition	Pillai's Trace	.045	1.472	36.000	3474.000	.035	.015
	Wilks' Lambda	.956	1.471	36.000	3416.259	.035	.015
	Hotelling's Trace	.046	1.470	36.000	3464.000	.035	.015
	Roy's Largest Root	.019	1.842[b]	12.000	1158.000	.038	.019

a. Exact statistic
b. The statistic is an upper bound on F that yields a lower bound on the significance level.
c. Design: Intercept + NumericCondition

Table 7

Self-Efficacy—Descriptives

				95% Confidence Interval	
Dependent Variable	Condition	Mean	Std. Error	Lower Bound	Upper Bound
Answer.EfficacyQ1	Generic	4.057	.112	3.836	4.277
	Choice	4.066	.115	3.841	4.291
	Fantasy	3.954	.112	3.735	4.173
	Circuit	3.663	.117	3.434	3.892
Answer. EfficacyQ2	Generic	5.213	.098	5.022	5.405
	Choice	5.267	.100	5.072	5.463
	Fantasy	5.273	.097	5.083	5.463
	Circuit	5.000	.101	4.801	5.199
Answer. EfficacyQ3	Generic	5.267	.103	5.064	5.469
	Choice	5.392	.105	5.186	5.599
	Fantasy	5.431	.103	5.230	5.632
	Circuit	5.108	.107	4.898	5.318
Answer. EfficacyQ4	Generic	3.300	.108	3.089	3.511
	Choice	3.042	.110	2.826	3.257
	Fantasy	3.076	.107	2.866	3.286
	Circuit	2.824	.112	2.605	3.044
Answer. EfficacyQ5	Generic	4.220	.110	4.005	4.435
	Choice	4.090	.112	3.871	4.310
	Fantasy	3.918	.109	3.704	4.131
	Circuit	3.932	.114	3.709	4.155
Answer. EfficacyQ6	Generic	4.303	.106	4.096	4.511
	Choice	4.132	.108	3.920	4.344
	Fantasy	3.980	.105	3.774	4.186
	Circuit	3.932	.110	3.717	4.147
Answer. EfficacyQ7	Generic	3.973	.099	3.779	4.168
	Choice	3.799	.101	3.600	3.997
	Fantasy	3.822	.098	3.629	4.015
	Circuit	3.559	.103	3.358	3.760
Answer. EfficacyQ8	Generic	4.060	.099	3.866	4.254
	Choice	3.851	.101	3.652	4.049
	Fantasy	3.789	.098	3.596	3.983
	Circuit	3.563	.103	3.361	3.764
Answer. EfficacyQ9	Generic	4.433	.107	4.224	4.643
	Choice	4.295	.109	4.082	4.509
	Fantasy	4.263	.106	4.055	4.471
	Circuit	4.222	.111	4.005	4.439
Answer. EfficacyQ10	Generic	4.790	.104	4.587	4.993
	Choice	4.462	.106	4.254	4.669
	Fantasy	4.599	.103	4.397	4.801
	Circuit	4.419	.107	4.209	4.630
Answer. EfficacyQ11	Generic	5.677	.109	5.463	5.890
	Choice	5.552	.111	5.334	5.770
	Fantasy	5.589	.108	5.377	5.801
	Circuit	5.233	.113	5.012	5.454
Answer. EfficacyQ12	Generic	4.707	.108	4.495	4.918
	Choice	4.469	.110	4.253	4.685
	Fantasy	4.480	.107	4.270	4.691
	Circuit	4.172	.112	3.953	4.392

Table 8

Self-Efficacy—Posthocs

Dependent Variable	Conditions	p-value
Complete a program if I had no help at all	Generic > Circuit	p < .05
Complete a program if I had no help at all	Choice > Circuit	p < .05
Write a program for an extremely difficult problem	Generic > Circuit	p < .05
Come up with a suitable strategy…	Generic > Circuit	p < .005
Manage my time efficiently…	Generic > Circuit	p < .05
Manage my time efficiently…	Choice > Circuit	p < .05
Write a program for a simple problem	Generic > Circuit	p < .005
Write a program for a simple problem	Choice > Circuit	p < .05
Write a program for a simple problem	Fantasy > Circuit	p < .05
Write a program for a moderately difficult problem	Generic > Circuit	p < .001
Write a program for a moderately difficult problem	Fantasy > Circuit	p < .05
Mentally trace through the execution of a long…	Generic > Circuit	p < .05
Mentally trace through the execution of a long…	Generic > Fantasy	p < .05
Write logically correct blocks of code	Generic > Choice	p < .05
Write logically correct blocks of code	Generic > Circuit	p < .05

Post hoc comparisons (LSD) revealed that participants in both the generic and choice conditions scored higher on “Complete a program if I had no help at all” than participants in the circuit condition. $p<.05$. Participants in the generic condition scored higher on “Write a program for an extremely difficult problem” than participants in the circuit condition, $p<.05$. Participants in the generic condition scored higher on “Come up with a suitable strategy for a given problem in a short time” than participants in the circuit condition, $p<.005$. Participants in both the generic and choice condition scored higher on “Manage my time efficiently if I had a pressing deadline on a problem” than participants in the circuit condition, $p<.05$. Participants in the generic, choice, and fantasy conditions scored higher on “Write a program for a simple problem” than participants in the circuit condition, $p<.05$. Participants in both the generic and fantasy condition scored higher on “Write a program for a moderately difficult problem” than participants in the circuit condition, $p<.05$. Participants in the generic condition scored higher on “Mentally trace through the execution of a long, complex, program given to me” than both the circuit and fantasy conditions, $p<.05$. Participants in the generic condition scored higher on “Write logically correct blocks of code” than participants in both the choice and circuit conditions, $p<.05$.

Table 9

GEQ—MANOVA Multivariate F-tests

Effect		Value	*F*	Hypothesis *df*	Error *df*	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.989	2477.999[a]	42.000	1126.000	.000	.989
	Wilks' Lambda	.011	2477.999[a]	42.000	1126.000	.000	.989
	Hotelling's Trace	92.430	2477.999[a]	42.000	1126.000	.000	.989
	Roy's Largest Root	92.430	2477.999[a]	42.000	1126.000	.000	.989
NumericCondition	Pillai's Trace	.233	2.265	126.000	3384.000	.000	.078
	Wilks' Lambda	.783	2.284	126.000	3374.672	.000	.078
	Hotelling's Trace	.258	2.303	126.000	3374.000	.000	.079
	Roy's Largest Root	.142	3.824[b]	42.000	1128.000	.000	.125

a. Exact statistic
b. The statistic is an upper bound on F that yields a lower bound on the significance level.
c. Design: Intercept + NumericCondition

Table 10

GEQ—Descriptives

				95% Confidence Interval	
Dependent Variable	Condition	Mean	Std. Error	Lower Bound	Upper Bound
Answer.g7flow	Generic	3.293	.064	3.167	3.420
	Choice	3.267	.066	3.138	3.396
	Fantasy	3.260	.064	3.134	3.385
	Circuit	3.326	.067	3.195	3.457
Answer.h8flow	Generic	2.650	.071	2.510	2.790
	Choice	2.760	.073	2.617	2.903
	Fantasy	2.582	.071	2.443	2.722
	Circuit	2.892	.074	2.747	3.038
Answer.i9flow	Generic	2.710	.077	2.559	2.861
	Choice	2.781	.079	2.627	2.935
	Fantasy	2.740	.076	2.590	2.890
	Circuit	2.864	.080	2.707	3.020
Answer.j10flow	Generic	3.803	.061	3.684	3.923
	Choice	3.795	.062	3.674	3.917
	Fantasy	3.793	.060	3.674	3.911
	Circuit	3.871	.063	3.747	3.995
Answer.k11flow	Generic	2.377	.068	2.243	2.511
	Choice	2.403	.070	2.266	2.539
	Fantasy	2.303	.068	2.170	2.436
	Circuit	2.491	.071	2.352	2.630
Answer.l12flow	Generic	3.340	.069	3.205	3.475
	Choice	3.424	.070	3.286	3.561
	Fantasy	3.385	.068	3.251	3.519
	Circuit	3.459	.071	3.319	3.598
Answer.m13imm	Generic	2.093	.071	1.953	2.234
	Choice	2.066	.073	1.923	2.209
	Fantasy	2.039	.071	1.900	2.179
	Circuit	2.211	.074	2.066	2.357
Answer.n14imm	Generic	2.443	.061	2.323	2.564
	Choice	3.118	.063	2.995	3.241
	Fantasy	3.211	.061	3.091	3.330
	Circuit	3.086	.064	2.961	3.211
Answer.o15imm	Generic	2.833	.071	2.693	2.973
	Choice	2.927	.073	2.784	3.070
	Fantasy	3.069	.071	2.930	3.208
	Circuit	3.050	.074	2.905	3.195
Answer.p16imm	Generic	2.213	.073	2.071	2.356
	Choice	2.153	.074	2.007	2.298
	Fantasy	2.336	.072	2.194	2.477
	Circuit	2.534	.075	2.386	2.682
Answer.q17imm	Generic	2.780	.065	2.653	2.907
	Choice	2.983	.066	2.853	3.112
	Fantasy	2.993	.064	2.867	3.120
	Circuit	3.082	.067	2.951	3.214
Answer.r18imm	Generic	2.473	.066	2.344	2.603
	Choice	2.649	.068	2.517	2.782
	Fantasy	2.661	.066	2.532	2.790
	Circuit	2.749	.069	2.614	2.884
Answer.s19comp	Generic	3.187	.071	3.048	3.326
	Choice	3.247	.072	3.105	3.388

	Fantasy	3.257	.070	3.118	3.395
	Circuit	3.115	.074	2.970	3.259
Answer.t20comp	Generic	2.170	.067	2.039	2.301
	Choice	2.375	.068	2.242	2.508
	Fantasy	2.362	.066	2.232	2.492
	Circuit	2.280	.069	2.144	2.415
Answer.u21comp	Generic	3.147	.065	3.019	3.274
	Choice	3.181	.066	3.051	3.310
	Fantasy	3.118	.064	2.992	3.245
	Circuit	2.867	.067	2.735	2.999
Answer.v22comp	Generic	3.227	.069	3.092	3.362
	Choice	3.226	.070	3.088	3.363
	Fantasy	3.211	.068	3.077	3.345
	Circuit	3.068	.071	2.928	3.208
Answer.w23comp	Generic	3.020	.064	2.894	3.146
	Choice	2.917	.066	2.788	3.045
	Fantasy	3.013	.064	2.888	3.138
	Circuit	2.649	.067	2.518	2.779
Answer.x24comp	Generic	3.317	.066	3.188	3.446
	Choice	3.309	.067	3.177	3.441
	Fantasy	3.375	.065	3.247	3.503
	Circuit	2.968	.068	2.834	3.102
Answer.y25chal	Generic	3.510	.069	3.375	3.645
	Choice	3.576	.070	3.439	3.714
	Fantasy	3.484	.068	3.349	3.618
	Circuit	3.613	.071	3.473	3.753
Answer.z26chal	Generic	2.980	.063	2.856	3.104
	Choice	3.003	.065	2.877	3.130
	Fantasy	3.072	.063	2.949	3.196
	Circuit	3.355	.066	3.226	3.484
Answer.za27chal	Generic	3.330	.065	3.202	3.458
	Choice	3.413	.067	3.282	3.544
	Fantasy	3.359	.065	3.231	3.486
	Circuit	3.487	.068	3.355	3.620
Answer.zb28chal	Generic	3.657	.062	3.535	3.778
	Choice	3.809	.063	3.685	3.933
	Fantasy	3.780	.062	3.659	3.900
	Circuit	3.961	.064	3.835	4.087
Answer.zc29chal	Generic	3.193	.065	3.066	3.321
	Choice	3.115	.066	2.984	3.245
	Fantasy	3.250	.065	3.123	3.377
	Circuit	3.423	.068	3.290	3.555
Answer.zd30chal	Generic	1.733	.063	1.610	1.856
	Choice	1.701	.064	1.576	1.827
	Fantasy	1.687	.062	1.565	1.810
	Circuit	1.943	.065	1.815	2.070
Answer.ze31tens	Generic	1.943	.065	1.816	2.070
	Choice	1.965	.066	1.836	2.095
	Fantasy	1.947	.064	1.821	2.073
	Circuit	2.262	.067	2.130	2.393
Answer.zf32tens	Generic	1.843	.059	1.727	1.960
	Choice	1.736	.061	1.617	1.855
	Fantasy	1.773	.059	1.657	1.889
	Circuit	1.996	.062	1.875	2.117
Answer.zg33tens	Generic	2.360	.074	2.215	2.505
	Choice	2.281	.076	2.133	2.430
	Fantasy	2.286	.074	2.142	2.431
	Circuit	2.520	.077	2.369	2.671

Answer.zh34tens	Generic	1.933	.068	1.800	2.067
	Choice	1.903	.069	1.766	2.039
	Fantasy	1.812	.068	1.680	1.945
	Circuit	2.090	.071	1.951	2.228
Answer.zi35tens	Generic	2.350	.076	2.200	2.500
	Choice	2.351	.078	2.198	2.503
	Fantasy	2.224	.076	2.075	2.372
	Circuit	2.545	.079	2.390	2.700
Answer.zj36tens	Generic	1.707	.060	1.589	1.824
	Choice	1.618	.061	1.498	1.738
	Fantasy	1.599	.059	1.482	1.715
	Circuit	1.914	.062	1.792	2.036
Answer.zk37pos	Generic	2.743	.061	2.625	2.862
	Choice	2.844	.062	2.723	2.965
	Fantasy	2.842	.060	2.724	2.960
	Circuit	2.742	.063	2.619	2.865
Answer.zl38pos	Generic	2.433	.069	2.299	2.568
	Choice	2.597	.070	2.460	2.735
	Fantasy	2.589	.068	2.455	2.722
	Circuit	2.448	.071	2.309	2.588
Answer.zm39pos	Generic	2.800	.062	2.679	2.921
	Choice	2.903	.063	2.779	3.026
	Fantasy	2.872	.061	2.752	2.992
	Circuit	2.889	.064	2.763	3.014
Answer.zn40pos	Generic	2.983	.061	2.864	3.103
	Choice	3.108	.062	2.985	3.230
	Fantasy	3.109	.061	2.990	3.228
	Circuit	3.079	.063	2.955	3.203
Answer.zo41pos	Generic	3.417	.065	3.290	3.543
	Choice	3.465	.066	3.336	3.594
	Fantasy	3.418	.064	3.292	3.544
	Circuit	3.444	.067	3.313	3.576
Answer.zp42pos	Generic	3.357	.067	3.225	3.489
	Choice	3.399	.069	3.264	3.534
	Fantasy	3.395	.067	3.264	3.526
	Circuit	3.387	.070	3.250	3.524
Answer.zq43neg	Generic	2.537	.065	2.409	2.664
	Choice	2.580	.066	2.449	2.710
	Fantasy	2.618	.065	2.491	2.745
	Circuit	2.437	.068	2.305	2.570
Answer.zr44neg	Generic	2.457	.071	2.317	2.596
	Choice	2.396	.073	2.253	2.538
	Fantasy	2.444	.071	2.305	2.583
	Circuit	2.538	.074	2.393	2.682
Answer.zs45neg	Generic	2.283	.071	2.144	2.422
	Choice	2.240	.072	2.098	2.381
	Fantasy	2.247	.070	2.109	2.385
	Circuit	2.308	.073	2.164	2.452
Answer.zt46neg	Generic	1.937	.058	1.822	2.051
	Choice	1.833	.059	1.717	1.950
	Fantasy	1.974	.058	1.860	2.087
	Circuit	1.896	.060	1.777	2.015
Answer.zu47neg	Generic	2.930	.081	2.771	3.089
	Choice	3.003	.083	2.841	3.165
	Fantasy	3.105	.080	2.948	3.263
	Circuit	2.910	.084	2.746	3.075
Answer.zv48neg	Generic	1.610	.055	1.503	1.717
	Choice	1.500	.056	1.391	1.609

	Fantasy	1.513	.054	1.407	1.619
	Circuit	1.699	.057	1.588	1.810

Table 11

GEQ—Posthocs

Dependent Variable	Conditions	p-value
"I forgot everything around me" (flow)	Circuit > Generic, Fantasy	p < .05
"It was aesthetically pleasing" (immersion)	Circuit, Fantasy, Choice > Generic	p < .001
"I felt that I could explore things" (immersion)	Circuit > Generic, Choice	p < .005
"I found it impressive" (immersion)	Circuit, Fantasy, Choice > Generic	p < .05
"It felt like a rich experience" (immersion)	Circuit, Fantasy > Generic	p < .05
"I was good at it" (competence)	Generic, Choice, Fantasy > Circuit	p < .01
"I was fast at reaching the game's targets" (competence)	Generic, Choice, Fantasy > Circuit	p < .005
"I felt competent" (competence)	Generic, Choice, Fantasy > Circuit	p < .001
"I thought it was hard" (challenge)	Circuit > Generic, Choice, Fantasy	p < .005
"I felt challenged" (challenge)	Circuit > Generic, Fantasy	p < .05
"I had to put a lot of effort into it" (challenge)	Circuit > Generic, Choice	p < .05
"I felt time pressure" (challenge)	Circuit > Generic, Choice, Fantasy	p < .05
"I felt tense" (tension)	Circuit > Generic, Choice, Fantasy	p < .005
"I felt restless" (tension)	Circuit > Choice, Fantasy	p < .01
"I felt irritable" (tension)	Circuit > Fantasy	p < .005
"I felt frustrated" (tension)	Circuit > Fantasy	p < .005
"I felt pressured" (tension)	Circuit > Generic, Choice, Fantasy	p < .05
"I felt tense" (tension)	Circuit > Generic, Choice, Fantasy	p < .005
"I felt tense" (tension)	Circuit > Generic, Choice, Fantasy	p < .005
"I felt tense" (tension)	Circuit > Generic, Choice, Fantasy	p < .005
"It gave me a bad mood" (negative affect)	Circuit > Fantasy, Choice	p < .05

Post hoc comparisons (LSD) revealed that participants in the circuit condition scored higher on "I forgot everything around me" (flow) than participants in the generic and fantasy conditions, $p<.05$. Participants in the circuit, fantasy, and choice conditions scored higher on "It was aesthetically pleasing" (immersion) than participants in the generic condition, $p<.001$. Participants in the circuit condition scored higher on "I felt that I could explore things" (immersion) than participants in the generic and choice conditions, $p<.005$. Participants in the circuit, fantasy, and choice conditions scored higher on "I found it impressive" (immersion) than participants in the generic condition, $p<.05$. Participants in the circuit and fantasy conditions scored higher on "It felt like a rich experience" (immersion) than participants in the generic condition, $p<.05$. Participants in the generic, choice, and fantasy conditions scored higher on "I was good at it" (competence) than participants in the circuit condition, $p<.01$. Participants in the generic, choice, and fantasy conditions scored higher on "I was fast at reaching the game's targets" (competence) than participants in the circuit condition, $p<.005$. Participants in the generic, choice, and fantasy conditions scored higher on "I felt competent" (competence) than participants in the circuit condition, $p<.001$. Participants in the circuit condition scored higher on "I thought it was hard" (challenge) than participants in the generic, choice, and fantasy conditions, $p<.005$. Participants in the circuit condition scored higher on "I felt challenged" (challenge) than

participants in the generic and fantasy conditions, $p<.05$. Participants in the circuit condition scored higher on "I had to put a lot of effort into it" (challenge) than participants in the generic and choice conditions, $p<.05$. Participants in the circuit condition scored higher on "I felt time pressure" (challenge) than participants in the generic, choice, and fantasy conditions, $p<.05$. Participants in the circuit condition scored higher on "I felt tense" (tension) than participants in the generic, choice, and fantasy conditions, $p<.005$. Participants in the circuit condition scored higher on "I felt restless" (tension) than participants in the choice and fantasy conditions, $p<.01$. Participants in the circuit condition scored higher on "I felt irritable" (tension) than participants in the fantasy condition, $p<.005$. Participants in the circuit condition scored higher on "I felt frustrated" (tension) than participants in the fantasy condition, $p<.005$. Participants in the circuit condition scored higher on "I felt pressured" (tension) than participants in the generic, choice, and fantasy conditions, $p<.05$. Participants in the circuit condition scored higher on "It gave me a bad mood" (negative affect) than participants in the fantasy and choice conditions, $p<.05$.

Table 12

PENS—MANOVA Multivariate F-tests

Effect		Value	*F*	Hypothesis *df*	Error *df*	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.957	1226.218[a]	21.000	1147.000	.000	.957
	Wilks' Lambda	.043	1226.218[a]	21.000	1147.000	.000	.957
	Hotelling's Trace	22.450	1226.218[a]	21.000	1147.000	.000	.957
	Roy's Largest Root	22.450	1226.218[a]	21.000	1147.000	.000	.957
NumericCondition	Pillai's Trace	.097	1.821	63.000	3447.000	.000	.032
	Wilks' Lambda	.905	1.840	63.000	3424.616	.000	.033
	Hotelling's Trace	.102	1.860	63.000	3437.000	.000	.033
	Roy's Largest Root	.075	4.104[b]	21.000	1149.000	.000	.070

a. Exact statistic
b. The statistic is an upper bound on F that yields a lower bound on the significance level.
c. Design: Intercept + NumericCondition

Table 13

PENS—Descriptives

Dependent Variable	Condition	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Answer.zw49penscomp	Generic	4.540	.098	4.348	4.732
	Choice	4.486	.100	4.290	4.682
	Fantasy	4.503	.097	4.312	4.694
	Circuit	4.025	.102	3.826	4.225
Answer.zx50penscomp	Generic	4.413	.095	4.227	4.600
	Choice	4.392	.097	4.202	4.582
	Fantasy	4.312	.094	4.128	4.497
	Circuit	3.982	.098	3.789	4.175
Answer.zy51penscomp	Generic	4.437	.099	4.243	4.631
	Choice	4.507	.101	4.309	4.705
	Fantasy	4.444	.098	4.252	4.637
	Circuit	4.065	.102	3.864	4.266
Answer.zz52pensauton	Generic	3.903	.106	3.696	4.111
	Choice	3.986	.108	3.774	4.198
	Fantasy	4.049	.105	3.843	4.256
	Circuit	4.068	.110	3.853	4.283
Answer.zza53pensauton	Generic	3.723	.104	3.519	3.928
	Choice	3.941	.106	3.732	4.150
	Fantasy	3.862	.104	3.659	4.065
	Circuit	4.014	.108	3.802	4.226
Answer.zzb54pensauton	Generic	2.633	.093	2.451	2.816
	Choice	2.778	.095	2.591	2.964
	Fantasy	2.852	.092	2.671	3.033
	Circuit	3.186	.096	2.997	3.376
Answer.zzc55pensrelatedness	Generic	1.970	.082	1.810	2.130
	Choice	2.017	.083	1.854	2.181
	Fantasy	2.053	.081	1.893	2.212
	Circuit	2.111	.085	1.945	2.277
Answer.zzd56pensrelatedness	Generic	1.893	.080	1.737	2.050
	Choice	1.882	.082	1.722	2.042
	Fantasy	1.928	.079	1.772	2.083
	Circuit	2.093	.083	1.931	2.256
Answer.zze57pensrelatedness_rev	Generic	5.283	.130	5.028	5.539
	Choice	5.406	.133	5.145	5.667
	Fantasy	5.273	.130	5.019	5.527
	Circuit	5.057	.135	4.792	5.323
Answer.zzf58penspresence	Generic	2.127	.088	1.955	2.299
	Choice	2.170	.090	1.995	2.346
	Fantasy	2.296	.087	2.125	2.467
	Circuit	2.387	.091	2.209	2.566
Answer.zzg59penspresence	Generic	1.793	.080	1.637	1.950
	Choice	1.740	.081	1.580	1.899
	Fantasy	1.951	.079	1.795	2.106
	Circuit	2.054	.083	1.892	2.216
Answer.zzh60penspresence	Generic	1.733	.077	1.583	1.884
	Choice	1.771	.078	1.617	1.924
	Fantasy	1.855	.076	1.706	2.005
	Circuit	1.968	.080	1.812	2.124

Answer.zzi61penspresence_rev	Generic	5.073	.131	4.816	5.331
	Choice	4.858	.134	4.595	5.121
	Fantasy	4.914	.131	4.658	5.171
	Circuit	4.645	.136	4.378	4.913
Answer.zzj62penspresence	Generic	2.053	.083	1.891	2.215
	Choice	2.059	.084	1.894	2.224
	Fantasy	2.007	.082	1.846	2.167
	Circuit	2.201	.086	2.033	2.369
Answer.zzk63penspresence	Generic	1.517	.070	1.379	1.654
	Choice	1.601	.072	1.460	1.741
	Fantasy	1.553	.070	1.416	1.689
	Circuit	1.601	.072	1.460	1.741
Answer.zzl64penspresence	Generic	1.767	.078	1.614	1.919
	Choice	1.771	.079	1.615	1.927
	Fantasy	1.865	.077	1.714	2.017
	Circuit	1.814	.081	1.655	1.972
Answer.zzm65penspresence	Generic	3.873	.106	3.665	4.082
	Choice	3.969	.109	3.756	4.182
	Fantasy	3.980	.106	3.773	4.188
	Circuit	4.115	.110	3.898	4.331
Answer.zzn66penspresence	Generic	1.607	.068	1.472	1.741
	Choice	1.663	.070	1.526	1.800
	Fantasy	1.674	.068	1.541	1.808
	Circuit	1.796	.071	1.656	1.935
Answer.zzo67penscontrols	Generic	4.773	.102	4.574	4.973
	Choice	4.580	.104	4.376	4.784
	Fantasy	4.664	.101	4.466	4.863
	Circuit	4.699	.106	4.492	4.906
Answer.zzp68penscontrols	Generic	3.950	.104	3.746	4.154
	Choice	3.743	.106	3.535	3.951
	Fantasy	3.937	.103	3.735	4.140
	Circuit	4.032	.108	3.821	4.243
Answer.zzq69penscontrols	Generic	4.723	.103	4.521	4.926
	Choice	4.437	.105	4.231	4.644
	Fantasy	4.681	.103	4.480	4.882
	Circuit	4.613	.107	4.403	4.823

Table 14

PENS—Posthocs

Dependent Variable	Conditions	p-value
PENS_competence_1	Circuit < Generic, Choice, Fantasy	p < .001
PENS_competence_2	Circuit < Generic, Choice, Fantasy	p < .05
PENS_competence_3	Circuit < Generic, Choice, Fantasy	p < .01
PENS_autonomy_3	Circuit > Generic, Choice, Fantasy	p < .05
PENS_presence_2	Circuit > Generic, Choice	p < .05
PENS_presence_6	Circuit > Generic, Choice, Fantasy	p < .005

Across all 3 questions on competence, participants in the circuit condition scored lower than participants in the generic, choice, and fantasy conditions, $p<.05$. For the question on autonomy, participants in the circuit condition scored higher than participants in the generic, choice, and fantasy conditions, $p<.05$. For one of the two questions on presence, participants in the circuit condition scored higher than participants in the generic, and choice conditions, $p<.05$. For the other of the two questions on presence, participants in the circuit condition scored higher than participants in the generic, choice and fantasy conditions, $p<.005$.

Table 15

Choice Condition—Descriptives

Dependent Variable	Choice	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Levels Completed	Generic	7.959	.374	7.222	8.696
	Fantasy	8.135	.263	7.618	8.653
	Circuit	7.791	.391	7.022	8.560
Hints Requested	Generic	1.699	.358	.995	2.402
	Fantasy	2.608	.251	2.114	3.102
	Circuit	2.925	.373	2.191	3.660
Attempts	Generic	21.562	4.124	13.444	29.679
	Fantasy	33.034	2.896	27.333	38.735
	Circuit	41.836	4.305	33.363	50.309

References

Asgari, M., & Kaufman, D. (2004). Relationships Among Computer Games, Fantasy, and Learning. *Proceedings of the 2nd International Conference on Imagination and Education*, 1–8. doi:10.1007/s13398-014-0173-7.2

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *The British Journal of Educational Psychology*, *72*(Pt 2), 261–278. doi:10.1348/000709902158883

Barab, S., & Duffy, T. (2000). From practice fields to communities of practice. *Theoretical Foundations of Learning Environments*, (1), 25–56. doi:10.1037/0022-3514.90.4.644

Becker, D. A. (1997). The Effects Of Choice On Auditors' Intrinsic Motivation and Performance. *Behavioral Research in Accounting*, *9*.

Betz, N. E., & Hackett, G. (1986). Applications of Self-Efficacy Theory to Understanding Career Choice Behavior. *Journal of Social and Clinical Psychology*, *4*(3), 279–289. doi:10.1521/jscp.1986.4.3.279

Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2005). Motivation and Cognitive Engagement in Learning Environments. In *The Cambridge Handbook of the Learning Sciences* (pp. 475–488). doi:http://dx.doi.org/10.1017/CBO9780511816833.029

Brauner, P., Leonhardt, T., Ziefle, M., & Schroeder, U. (2010). The effect of tangible artifacts, gender and subjective technical competence on teaching programming to seventh graders. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5941 LNCS, pp. 61–71). doi:10.1007/978-3-642-11376-5_7

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*(1), 182–197. doi:10.1037/0022-0663.98.1.182

Clark, R. C., & Mayer, R. E. (2011). E-learning and the Science of Instruction : Proven Guidelines for Consumers and Designers of Multimedia Learning.

Classcraft. (2016). Retrieved from http://classcraft.com/

CodeCombat. (2016). Retrieved from https://codecombat.com/

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*(4), 715–730. doi:10.1037/0022-0663.88.4.715

Crowley, K., & Jacobs, M. (2002). Building Islands of Expertise in Everyday Family Activity. *Learning Conversations in Museums*, 1–23.

Daanen, H., & Grant, L. (2007). Space Mission: Ice Moon. In *ACM SIGGRAPH 2007 educators program on - SIGGRAPH '07* (p. 19). New York, New York, USA: ACM Press. doi:10.1145/1282040.1282060

Deci, E., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press.

Dewey, J. (1913). Interest and effort in education. *Search*, 1859–1952.

Driskell, J. E., & Dwyer, D. J. (1984). Microcomputer Videogame Based Training. *Educational Technology*, *24*, 11–17.

Entertainment Software Assotiation. (2015). Sales, Demographic, and Usage Data, *4*(1), 2–4. Retrieved from http://www.theesa.com/facts/pdfs/ESA_EF_2008.pdf

Eshel, Y., & Kohavi, R. (2003). Perceived Classroom Control, Self-Regulated Learning Strategies, and Academic Achievement. *Educational Psychology*, *23*(3), 249. doi:10.1080/0144341032000060093

Evans, M., & Boucher, A. R. (2015). Optimizing the power of choice: Supporting student autonomy to foster motivation and engagement in learning. *Mind, Brain, and Education*, *9*(2), 87–91. doi:10.1111/mbe.12073

Flowerday, T., & Schraw, G. (2000). Teacher beliefs about instructional choice: A phenomenological study. *Journal of Educational Psychology*, *92*(4), 634–645. doi:10.1037/0022-0663.92.4.634

Fullerton, T. (2014). What games do well. *Postsecondary Play: The Role of Games and Social Media in Higher Education*.

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of "Seductive Details" on Macroprocessing and Microprocessing in Adults and Children. *Source: Cognition and Instruction*, *6*(1), 41–57. doi:10.1207/s1532690xci0601_2

Garris, R., & Ahlers, R. (2001). A Game-Based Training Model: Development, Application, And Evaluation. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*. Retrieved from http://ntsa.metapress.com/app/home/contribution.asp?referrer=parent&backto=issue,6,151;journal,6,7;linkingpublicationresults,1:113340,1

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming*, *33*(4), 441–467. doi:10.1177/1046878102238607

Gee, J. P. (2008). Game-like learning: An example of situated learning and implications for opportunity to learn. *Assessment, Equity, and Opportunity to Learn*, 200–221. doi:http://dx.doi.org/10.1017/CBO9780511802157.010

Goetz, E. T., & Sadoski, M. (1995). Commentary: The Perils of Seduction: Distracting Details or Incomprehensible Abstractions? *Reading Research Quarterly*, *30*(3), 500–511. doi:10.2307/747628

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*(1), 5–26. doi:10.1037/0003-066X.53.1.5

Hardre, P. L., & Reeve, J. (2003). A motivational model of rural students' intentions to persist in, versus drop out of, high school. *Journal of Educational Psychology*, *95*(2), 347–356. doi:10.1037/0022-0663.95.2.347

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414–434. doi:10.1037/0022-0663.90.3.414

Harteveld, C., & Sutherland, S. (2015). The Goal of Scoring: Exploring the Role of Game Performance in Educational Games. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015)*.

Harrell, S. V., & Harrell, D. (2009). Exploring the Potential of Computational Self-Representations for Enabling Learning: Examining At-risk Youths' Development of Mathematical/Computational Agency. *Proceedings of the Digital Arts and Culture Conference*.

IJsselsteijn, W., De Kort, Y., Poels, K., Jurgelionis, A., & Bellotti, F. (2007). Characterising and Measuring User Experiences in Digital Games. *International Conference on Advances in Computer Entertainment Technology*, *620*, 1–4. doi:10.1007/978-1-60761-580-4

Jolivette, K., Wehby, J. H., Canale, J., & Massey, N. G. (2001). Effect of choice-making opportunities on the behavior of students with emotional and behavioral disorders. *Behavioral Disorders*, *26*(2), 131–145.

Kao, D., & Harrell, D. F. (2015a). Exploring construction, play, use of virtual identities in STEM learning. *Jean Piaget Society Annual Conference*.

Kao, D., & Harrell, D. F. (2015b). Exploring the Impact of Role Model Avatars on Game Exp. in Educational Games. *ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (CHI PLAY)*.

Kao, D., & Harrell, D. F. (2015c). Toward Avatar Models to Enhance Performance and Engagement in Educational Games. In *Computational Intelligence in Games*.

Kao, D., & Harrell, D. F. (2015d). Toward Evaluating the Impacts of Virtual Identities on STEM Learning. *Foundations of Digital Games*.

Kao, D., & Harrell, D. F. (2015e). Mazzy: A STEM Learning Game. *Foundations of Digital Games*.

Kao, D., & Harrell, D. F. (2016a). Toward Understanding the Impacts of Role Model Avatars on Engagement in Computer Science Learning. *The Annual Meeting of the American Educational Research Association (AERA)*.

Kao, D., & Harrell, D. F. (2016b). Exploring the Effects of Encouragement in Educational Games. *ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI 2016)*.

Kao, D., & Harrell, D. F. (2016c). Exploring the Impact of Avatar Color on Game Experience in Educational Games. *ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI 2016)*.

Kao, D., & Harrell, D. F. (2016d). Exploring the Effects of Dynamic Avatars on Performance and Engagement in Educational Games. In *Games+Learning+Society (GLS 2016)*.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review*, *19*(4), 429–442. doi:10.1007/s10648-006-9027-y

Kernan, M., Heimann, B., & Hanges, P. (1991). Effects of goal choice, strategy choice, and feedback source on goal acceptance, performance, and subsequent goals. *Journal of Applied Social Psychology*, *21*, 713–733.

Lave, J. (1991). Situating learning in communities of practice. *Perspectives on Socially Shared Cognition*, 63–82. doi:10.1037/10096-003

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. *Learning in Doing*, *95*, 138. doi:10.2307/2804509

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, *32*(4), 569–587. doi:10.1016/j.cedpsych.2006.07.002

Lemke, J. (1997). Cognition, context, and learning: A social semiotic perspective. *Situated Cognition: Social, Semiotic, and Psychological Perspectives*, (1972), 37–55. Retrieved from http://books.google.com/books?id=ivqlQw9Dy-8C&pgis=1

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. *Aptitude Learning and Instruction*. doi:10.1016/S0037-6337(09)70509-1

Martin, J. E., Mithaug, D. E., Cox, P., Peterson, L. Y., Van Dycke, J. L., & Cash, M. E. (2003). Increasing self-determination: Teaching students to plan, work, evaluate, and adjust. *Exceptional Children*, *69*(4), 431–447. doi:doi: 10.1177/001440290306900403

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. *IThe Cambridge Handbook of Multimedia Learning* (Vol. 16). doi:10.1075/idj.16.1.13pel

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology. Applied*, *11*(4), 256–265. doi:10.1037/1076-898X.11.4.256

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*(1), 187–198. doi:10.1037/0022-0663.93.1.187

Ozdemir, D., & Doolittle, P. (2015). Revisiting the Seductive Details Effect in Multimedia Learning: Context-Dependency of Seductive Details. *Journal of Educational Multimedia and Hypermedia*, *24*(2), 101–119.

Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings. *Review of Educational Research*, *66*(4), 543–578. doi:10.3102/00346543066004543

Park, B., Flowerday, T., & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, *44*, 267–278. doi:10.1016/j.chb.2014.10.061

Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*, *27*(1), 5–10. doi:10.1016/j.chb.2010.05.006

Parker, L. E., & Lepper, M. R. (1992). Effects of Fantasy Contexts on Children's Learning and Motivation. *Journal of Personality and Social Psychology*, *62*(4), 625–633. doi:10.1037/0022-3514.62.4.625

Patall, E. a, Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: a meta-analysis of research findings. *Psychological Bulletin*, *134*(2), 270–300. doi:10.1037/0033-2909.134.2.270

Ramalingam, V., & Wiedenbeck, S. (1998). Development and Validation of Scores on a Computer Programming Self-Efficacy Scale and Group Analyses of Novice Programmer Self-Efficacy. *Journal of Educational Computing Research*, *19*(4), 367–381. doi:10.2190/C670-Y3C8-LTJ1-CT3P

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, *16*(9), 13–20. doi:10.3102/0013189X029002004

Resnick, M., Flanagan, M., Kelleher, C., MacLaurin, M., Ohshima, Y., Perlin, K., & Torres, R. (2009). Growing Up Programming Democratizing the Creation of Dynamic, Interactive Media. *CHI EA '09 Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems, April 4-9, 2009, Boston, Massachusetts, USA*, 3293–3296. doi:10.1145/1520340.1520472

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*. doi:10.1016/j.edurev.2012.05.003

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development*, *44*(2), 43–58. doi:10.1007/BF02300540

Rose, D. H., & Meyer, A. (2002). Teaching Every Student in the Digital Age: Universal Design for Learning. *Ericedgov*, 216. doi:10.1007/s11423-007-9056-3

Rowland, E., Skinner, C. H., Davis-Richards, K., Saudargas, R., & Robinson, D. H. (2008). An Investigation of Placement and Type of Seductive Details: The Primacy Effect of Seductive Details on Text Recall, *15*(2), 80–90.

Rummer, R., Schweppe, J., Fürstenberg, A., Seufert, T., & Brünken, R. (2010). Working memory interference during processing texts and pictures: Implications for the explanation of the modality effect. *Applied Cognitive Psychology*, *24*(2), 164–176. doi:10.1002/acp.1546

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68–78. doi:10.1037/0003-066X.55.1.68

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, *30*(4), 344–360. doi:10.1007/s11031-006-9051-8

Sanchez, C. a, & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, *34*(2), 344–55. doi:10.3758/BF03193412

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction*, *19*(6), 481–494. doi:10.1016/j.learninstruc.2008.08.001

Scott, M. J., & Ghinea, G. (2013). Integrating Fantasy Role-Play Into the Programming Lab. *Proceeding of the 44th ACM Technical Symposium on Computer Science Education - SIGCSE '13*, 119. doi:10.1145/2445196.2445237

Sfard, A. (1998). On Two Metaphors for learning and the Dangers of Choosing Just One. *Educational Researcher*, *27*(2), 4–13. doi:10.3102/0013189X027002004

Shaffer, D. W. (2006). Epistemic frames for epistemic games. *Computers & Education*, *46*(3), 223–234. doi:10.1016/j.compedu.2005.11.003

Thalheimer, W. (2004). Bells, whistles, neon, and purple prose: When interesting words, sounds, and visuals hurt learning and performance — a review of the seductive-augmentation research, 1–29.

Thoits, P. A. (1983). Multiple identities and psychological well-being: A reformulation and test of the social isolation hypothesis. *American Sociological Review*, *48*(2), 174–187. doi:10.2307/2095103

Virvou, M., Katsionis, G., & Manos, K. (2005). Combining software games with education: Evaluation of its educational effectiveness. *Educational Technology and Society*. doi:10.1016/j.corsci.2007.02.007

Walkerdine, V. (1997). Redefining the subject in situated cognition theory. In *Situated cognition : social, semiotic, and psychological perspectives* (pp. 57–70).

Wenger, E., Choir, A. S., & Fitriyah, R. (1998). Communities of Practice: Learning, Meaning, and Identity. *Systems Thinker*, *9*, 2–3. doi:10.1016/j.jchas.2013.03.426

Westrom, M., & Shaban, A. (1992). Intrinsic motivation in microcomputer games. *Journal of Research on Computing in Education*, *24*(4), 433. doi:10.1080/08886504.1992.10782018

Wilson, K. a., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., … Conkey, C. (2009). Relationships Between Game Attributes and Learning Outcomes: Review and Research Proposals. *Simulation & Gaming*, *40*(2), 217–266. doi:10.1177/1046878108321866

Young, M. F. (1993). Instructional design for situated learning. *Educational Technology Research and Development*, *41*(1), 43–58. doi:10.1007/BF02297091

Yuan, L., & Powell, S. (2013). MOOCs and Open Education: Implications for Higher Education. *Cetis*, 19. doi:http://publications.cetis.ac.uk/2013/667

Zimmerman, E. (2011). Let the Games Be Games: Aesthetics, Instrumentalization & Game Design. In *Presentation at Game Developers Conference*.