Using a Multi-Dimensional Model of Gender to Assess Learning with Different Game-Based Learning Narratives

Hayden Stec[1], J. Elizabeth Richey[2], Huy Nguyen[1], Nicole Else-Quest[3], Jessica Hammer[1], Ryan S. Baker[4], Ivon Arroyo[5], Bruce McLaren[1]

[1] Carnegie Mellon University, [2] University of Pittsburgh, [3] University of North Carolina at Chapel Hill, [4] University of Pennsylvania, [5] University of Massachusetts Amherst

Abstract. Digital learning games can help address gender disparities in math by promoting better learning experiences and outcomes for girls. However, there is a need for more research to understand why some digital learning games might be especially effective for girls studying mathematics. In this study, we assess two possible pathways: that girls might benefit from math games because they reduce the anxiety and evaluation apprehension that girls are more likely to experience when doing math; and that girls might benefit from math games when they enjoy the narrative and thus experience greater engagement. To evaluate these pathways, our work uses multiple dimensions of gender (e.g., gender identity and gender-typed interests, activities, and traits) and surveys of affective experiences to examine the impact of three learning systems with identical learning content: a digital learning game, *Decimal Point*, that has consistently led to better learning for girls over boys; a new masculine-typed game, *Ocean Adventure*, developed based on a survey of over 300 students; and a conventional tutoring system. We predicted that girls and students with stronger feminine-typed characteristics would experience less math anxiety in both *Decimal Point* and *Ocean Adventure* compared to the tutor. We also predicted that girls and students with stronger feminine-typed characteristics would experience greater engagement and learning with *Decimal Point* while boys and students with stronger masculine-typed characteristics would experience greater engagement and learning with *Ocean Adventure*. Consistent with predictions, students with stronger feminine-typed characteristics experienced less anxiety and evaluation apprehension in both games compared to the tutor. This suggests that math learning games may provide a way to address these negative affective experiences. In terms of our measures of engagement, we found that students with stronger masculine-typed characteristics reported greater experience of mastery in the masculine *Ocean Adventure*; however, this was the only indicator that the more masculine narrative of *Ocean Adventure* led to different experiences based on gender. This suggests that narrative alone may not have a strong enough effect on students based on gender, especially when other game features are kept constant. Contrary to our predictions, there were no effects of gender identity or condition on learning outcomes, although both masculine-typed and feminine-typed characteristics were negatively associated with learning. Overall, these results point to the value of a multi-dimensional model of gender in assessing learning with a game, the important role learning games can have in reducing math anxiety and evaluation apprehension for girls and students with feminine-typed characteristics, and the nuanced effects of game narratives on experiences with game-based learning.

Keywords: Game-based Learning, Decimals, Learning, Enjoyment, Gender, Anxiety, Evaluation Apprehension

Introduction

Digital learning games are instructional platforms that allow students to both engage playfully and learn (Gee, 2007; Mayer, 2014, 2019). In recent studies, some learning games have been shown to be more effective for girls than for boys, both in terms of learning and affective outcomes (Arroyo et al., 2013; McLaren et al., 2022; McLaren, Farzan, et al., 2017; Nguyen et al., 2022). These results suggest that games may be particularly effective for reducing gender disparities in math outcomes, including lower levels of motivation and math self-confidence among girls compared to boys (Else-Quest et al., 2010; Reilly et al., 2019). Inspired by these findings, researchers and designers have been interested in exploring how game-based learning can create more equitable learning experiences by promoting girls' math engagement and learning. However, their efforts often result in recommendations based on intuition rather than empirical evidence (Farrell & Moffat, 2014). An effective approach to addressing this issue requires a clear understanding of (1) how gender can be represented in an inclusive and meaningful way, as well as (2) how learning games influence the relationship between gender, affective experiences, and learning outcomes.

Our current work explores these questions through studies with the digital learning game *Decimal Point*, which teaches decimal numbers and operations to middle school students (McLaren, Adams, et al., 2017). While *Decimal Point* has been the subject of many classroom studies over the past 10 years, studies that have explored a variety of game-based learning questions, one highly consistent finding that emerged across studies is that girls have learned more than boys from the game after using *Decimal Point* to learn decimals (McLaren et al., 2022; McLaren, Farzan, et al., 2017; Nguyen et al., 2022).

Continuing this line of work, our current research investigates whether the narrative in a mathematics educational game may contribute to gender-based differences in learning, and whether learning math through games might reduce girls' math anxiety. For this study, we developed a new game, called *Ocean Adventure*, which features a narrative preferred by boys and students with masculine-typed characteristics but, at the same time, retains all learning activities and mathematics content from *Decimal Point*. Insights regarding gender preferences were derived from a survey in prior work investigating gender differences in game genre and narrative preferences (Nguyen et al., 2023).

We compared learning experiences and outcomes between the original *Decimal Point* game, the new narrative in *Ocean Adventure*, and a conventional tutor with identical instructional content but no game features. We have also utilized a multidimensional gender model, based on gender-typed occupational interests, activities, and traits (Liben et al., 2002), collectively referred to as "gender-typed characteristics", to explore students' learning experiences and outcomes using a more nuanced, continuous measure of gender.
With this study, we aim to test two possible explanations for the gender differences we have observed in learning outcomes. First, we test the possibility that girls are learning more because they enjoy the game narrative more than boys. We test this by measuring different aspects of enjoyment and by manipulating the narrative to be more appealing to boys versus girls based on the aforementioned game-based learning survey. Second, we test the possibility that girls are learning more because the game reduces their math anxiety and evaluation apprehension around math. Specifically, we test the following research questions and hypotheses:

- RQ1: Does the relationship between gender and learning outcomes differ depending on the game narrative?

Prior studies have shown that girls learn more from *Decimal Point* than boys (McLaren et al., 2022; McLaren, Farzan, et al., 2017; Nguyen et al., 2022), and we expect to replicate this effect. If girls are benefitting from the game because they enjoy the game narrative more than boys, then we predict that boys and students with more masculine-typed characteristics will learn more than girls and students with more feminine-typed characteristics in *Ocean Adventure*, which was specifically designed to be more appealing to boys. We hypothesize there will be no gender differences in learning outcomes in the tutor, which has no narrative and therefore should serve as a neutral condition.

- RQ2: Does the relationship between gender and enjoyment differ depending on the game narrative?

We specifically designed *Ocean Adventure* to be more appealing to boys based on game preference surveys we designed and administered (Nguyen et al., 2023). This survey found the proposed *Ocean Adventure* narrative to be more compelling to boys than girls, as well as identifying a number of game preferences between boys and girls that influenced the design of *Ocean Adventure*.

Therefore, we predict, first, that girls and students with more feminine-typed characteristics will report enjoying *Decimal Point* more than boys and students with more masculine-typed characteristics, and, second, that boys and students with more masculine-typed characteristics will report enjoying *Ocean Adventure* more than girls and students with more feminine-typed characteristics. We predict no gender or gender-typed characteristics differences in enjoyment in the tutor.

- RQ3: Do girls and students with more feminine-typed characteristics experience lower levels of anxiety when learning math through digital learning games compared to tutor instruction?

Prior research has shown that girls generally report greater levels of anxiety than boys when learning math and that this anxiety disrupts their math performance (Van Mier et al., 2019). We hypothesize that learning math through a digital learning game may be particularly beneficial for girls because it presents a less obvious math learning environment, which in turn reduces girls' thoughts of negative stereotypes about girls' math ability. Based on this, we predict that girls and students with more feminine-typed characteristics will report lower levels of anxiety in both game conditions compared to the tutor condition, while boys and students with more masculine-typed characteristics should report similar levels of anxiety across all conditions. Consistent with prior research, we predict that girls and students with more feminine-typed characteristics will express more anxiety than boys and students with more masculine-typed characteristics in the tutor, but they will have similar levels of anxiety in the games.

The examination of gender, learning experiences, and learning outcomes through our current research makes several important contributions to research on equitable game design. First, we report the impact of differing game narratives on students' learning experiences and outcomes. Through this experimental design, we test competing hypotheses about the reasons digital learning games might be especially beneficial to girls and students with more feminine-typed characteristics in math. Second, we identify the potential benefits in using a multidimensional model of gender to understand students' learning experiences in game environments. Based on these results, we highlight contributions to theories about game-based learning and provide insight on creating equitable learning experiences.

Background

Math, Game-Based Learning, and Gender

Prior research has identified several nuanced patterns of gender differences in mathematics education. While studies generally find minimal gender differences in math achievement (Ghasemi & Burley, 2019; Lindberg et al., 2010; Reilly et al., 2019), boys are still better represented among the top performers (Breda et al., 2018; Wai et al., 2010). More importantly, by late elementary school and persisting through the college years, girls and women tend to report less positive affect towards math (Levine & Pantoja, 2021), which in turn predicts their subsequent lack of STEM engagement and achievement (Else-Quest et al., 2013; Jiang et al., 2020). Additionally, girls tend to report more negative emotions and self-derogating attributions about their mathematics performance than boys (Mejía-Rodríguez et al., 2021), as well as higher levels of anxiety, which in turn hinder mathematics performance (Barroso et al., 2021; Van Mier et al., 2019).

Game-based learning could offer a series of benefits for girls in math education. Single-player and private digital games provide an environment for girls to practice mathematics without concern for external judgment from their peers and teachers (Niederle & Vesterlund, 2010), potentially reducing anxiety compared to more traditional math practice. Consistent with this idea, several studies testing the effects of STEM digital learning games have found gender differences favoring girls in experiences and outcomes (Joiner et al., 2011; Khan et al., 2017; Klisch et al., 2012; Tsai, 2017), although relatively few games other than *Decimal Point* have specifically examined gender and math learning (Nguyen et al., 2022). On the other hand, the differences in learning outcomes observed in *Decimal Point* might reflect gender differences in the appeal of the amusement park narrative, which was designed to be gender-neutral but has much in common with the puzzle-style design consistent with more girls' gaming preferences (Chou & Tsai, 2007). We tested these ideas in the current study by varying narratives and measuring affective experiences including math anxiety and enjoyment.

A recent synthesis by Hyde and colleagues (Hyde et al., 2019) has indicated that gender may be analyzed via multiple interrelated but separate dimensions. These dimensions can include birth-assigned gender, the most common operationalization of gender and following a binary categorization; gender identity, the internal sense of one's own gender; and gender-typed characteristics, the extent to which one aligns their behaviors, and preferences with stereotypical gender norms (Liben et al., 2002). Notably, prior work has shown that middle school children can be differentiated along these dimensions of gender even when their gender identity aligns with the traditional male and female categories (Cook et al., 2019).

The use of additional gender dimensions has also yielded benefits in game-based learning research, with a recent study showing that they model students' gaming preferences better than gender identity alone (Nguyen et al., 2023). Following this result, our work aimed to further investigate the connection between gender dimensions and learning outcomes, as well as enjoyment, from game play.

The Learning Game *Decimal Point* and the *Decimal Tutor*

Decimal Point is a digital learning game for mathematics which has been the subject of many classroom studies (Forlizzi et al., 2014; McLaren, 2024; McLaren, Adams, et al., 2017). As a web-based single player game, it can be easily accessed through school-issued computers or tablets. The game uses the theme of an amusement park to lead students through a variety of mini-games, each designed to teach students about decimals and decimal operations, with the content being targeted towards late elementary and early middle school students. In the base version of the game, the students progress through a linear path of theme areas (e.g. from Space Adventure to the Round Table to other theme areas), each with several mini-games (e.g. Escape the Aliens, The Joust) that address particular decimal topics and common misconceptions (Isotani et al., 2010). Students must complete every mini-game in this linear path in a set sequential order to traverse the entire amusement park and finish the game.

[Figure 1 about here]

Figure 1. The main game map in *Decimal Point* where students select mini-games to play.

Each mini-game features a problem-solving step, which belongs to one of five decimal problem types:

- sequence a given set of decimal numbers,
- complete a sequence of decimal numbers,
- put given decimals in less-than and greater-than buckets compared to a given single decimal,
- locate a given decimal number on the number line, and
- add two decimal numbers.

Each mini-game also includes a self-explanation step designed to reinforce student learning (Chi & Wylie, 2014). These activities are embedded in a playful narrative to promote both learning and enjoyment. For example, the mini-game Escape the Aliens places the player in a spaceship, running away from the "evil beast Xigorath" (Figure 2). Students see an arithmetic sequence of decimals on a locked door, with the first three numbers provided, and are prompted to find the next two in order to unlock the door. Although the individual mini-game has a tone of urgency and a sci-fi narrative, it is couched within *Decimal Point*'s lighthearted amusement park theme.

[Figure 2 about here]

Figure 2. The problem-solving activity (left) and self-explanation activity (right) in the Escape the Aliens mini-game.

The decimal tutor presents the mathematical content of *Decimal Point* using a more conventional user interface for tutoring systems and has been used in prior work with *Decimal Point* (McLaren, Adams, et al., 2017). Like *Decimal Point*, the tutor is accessed via an Internet browser for ease of use with school-issued laptops and tablets. In the tutor, the same problems as encountered in *Decimal Point* and *Ocean Adventure* are presented in the same order, but there is no game narrative, fantasy setting or game characters.

[Figure 3 about here]

Figure 3. The tutor's analogous problem to the Escape the Aliens mini-game, with the problem-solving activity (left) and self-explanation activity (right).

The New Learning Game *Ocean Adventure*

For this study, a new game, *Ocean Adventure*, was developed to explore the impact of differing game narratives on students' learning outcomes and enjoyment. *Ocean Adventure* features the same problems and problem order as *Decimal Point* and the tutor, but follows a different game narrative, in which seven pirate criminal masterminds are plotting to destroy the world with an atomic bomb placed on Doom Island. The player takes on the role of a special agent who confronts all the pirate leaders one by one in the game to eventually locate Doom Island and disarm the bomb. This narrative was chosen to align with the preferences of boys and those with strong masculine-typed characteristics, based on the findings from the previous survey of gaming preferences discussed earlier (Nguyen et al., 2023). The survey was conducted among 333 5th and 6th grade students; given four narratives to assess, including the *Ocean Adventure* narrative regarding an adventure at sea fighting pirates, 74% of boys but only 40% of girls found the proposed narrative compelling.

[Figure 4 about here]

Figure 4. The main game map in *Ocean Adventure* where students select mini-games to play.

In line with the narrative shift, the main game map has been transformed from an amusement park into a map marking pirate hideouts. Each mini-game was also redesigned to center around confronting various pirate crews, while retaining the same underlying mathematical content as *Decimal Point* and the tutor. For example, the mini-game "Cut through the Vines" involves the same decimal sequence to fill out as in the *Decimal Point* game and the tutor but introduces a new goal of cutting through layers of vines to infiltrate the ship of Hemlock, a pirate leader with plant-controlling powers.

[Figure 5 about here]

Figure 5. The “Cut through the Vines” mini-game, designed to be equivalent to the corresponding mini-game in *Decimal Point* (Figure 2) and problem in the tutor (Figure 3).

Method

Participants and Design

Four hundred and eighteen (418) students participated in the study, from six elementary and middle schools in a northeastern U.S. city. For a week of regularly scheduled class time, students completed the following materials in a self-paced sequence: a pretest, demographic survey, intervention materials, evaluation questionnaire, a posttest, and a week later, a delayed posttest. We randomly assigned each class to use either the games or the tutor; within the classes assigned to the game, each student was randomly assigned to play either *Decimal Point* or *Ocean Adventure*. This setting was chosen to avoid having students use both the games and the tutor in the same class, as it could be distracting for students to see their peers use a very different learning platform.

Seventy-nine (79) students were excluded from the analysis due to not finishing all of the study materials, with 25 assigned to play *Decimal Point*, 27 assigned to play *Ocean Adventure*, and 27 assigned to use the tutor. The age range of excluded students was from 10 to 12 years old (M = 10.58, SD = 0.65). Of the excluded students, 39% (n = 31) identified as male, 54% (n = 43) as female, and 6% (n = 5) did not disclose their gender.
Among the remaining 339 students, 122 played *Decimal Point*, 90 played *Ocean Adventure*, and 127 used the tutor. The age range of students in our final sample was from 10 to 12 years old (M = 10.75, SD = 0.63). In terms of gender, 50% (n = 169) of the students identified as male, 48% (n = 164) as female, 0.9% (n = 3) identified as non-binary, and 0.9% (n = 3) preferred not to disclose their gender. Due to the small sample size of the last two categories, they were not included in analyses of binary gender identity, but we still included them in analyses of gender-typed characteristics.

Materials

The primary materials used include *Decimal Point*, *Ocean Adventure* and the tutor. Additionally, three tests that were previously developed for *Decimal Point* and the tutor were also used, as well as questionnaires administered prior to and after the learning activity. All materials were developed in HTML/JavaScript and hosted on an Internet browser-based learner management system, which allowed students to access their assigned intervention and researchers to collect anonymized log data from their usage of the above materials.

Pretest, Posttest, and Delayed Posttest. There were three versions of the test with similar content, randomly assigned to each student as either the pretest, posttest, or delayed posttest. The tests consisted of 43 items; some items contained multiple components and were worth multiple points, and the total score available was 52 points. Test items assessed decimal skills and procedures which were directly addressed by the materials, or were conceptual questions not directly addressed. The three tests have been vetted through use on many other classroom studies (McLaren, 2024).

Demographic and Gender Surveys. A demographic survey was delivered prior to the pretest, and asked questions about age, grade level, gender identity, and race. This was followed by a 58-item Likert scale assessing gender-typed characteristics, leveraging prior work assessing these in relation to gender norms (Liben et al., 2002). The scale comprises the three domains of gender-typed occupational interests, activities, and traits, composed of items labeled as either masculine-typed or feminine-typed and rated from 1 (“not at all” / “never”) to 4 (“very much” / “very often”). The occupational interests domain contains 18 items referring to one’s interest in pursuing a given profession, such as “florist” (feminine) or “construction worker” (masculine). The activities domain contains 18 items measuring the frequency of engaging in a certain activity like “taking dance lessons” (feminine) or “going fishing” (masculine). The traits domain (22 items) evaluates students’ self-perceptions of qualities such as “neat” (feminine) or “adventurous” (masculine). In addition to these masculine-typed items and feminine-typed items, other options for students were made available in each domain, categorized as “neutral”, to accommodate students who do not identify with the masculine-typed or feminine-typed characteristics. Of note is that these gender-typed characteristics reflect stereotypical perceptions of gender differences, rather than true differences between boys and girls. That is, the scale measures the extent to which students’ behaviors align with conventional gender norms and stereotypes (Liben et al., 2002). To this end, two scales of feminine-typed characteristics ($\alpha = 0.81$) and masculine-typed characteristics ($\alpha = 0.83$) were generated by averaging the corresponding items from the occupational interests, activities, and traits domains. In sum, in addition to self-reported gender identity in the demographic survey, each student also had a score indicating their masculine-typed characteristics and feminine-typed characteristics, for a total of three gender dimensions.

Evaluation Questionnaire. After the intervention but prior to the posttest, students were given a questionnaire about their learning experience, rating statements on a Likert scale from 1 or “strongly disagree” to 5 or “strongly agree”. The questionnaire included seven constructs related to the students’ affective experience (Table 1): affective engagement (3 items, $\alpha = .71$ - (Ben-Eliyahu et al., 2018)); situational interest (3 items, $\alpha = 0.84$ - (Linnenbrink-Garcia et al., 2010)); enjoyment dimension of achievement emotion (6 items, $\alpha = 0.91$ - (Pekrun, 2005)); evaluation apprehension (4 items, $\alpha = 0.86$ - (Spencer et al., 1999)); state anxiety (3 items, $\alpha = 0.68$ - (Chung & Chang, 2017; Veit & Ware, 1983)).

Table 1. Example items for each affective experience construct measured in the evaluation survey.

Construct	Example statement
Affective engagement	I felt frustrated or annoyed.
Experience of meaning	The game [tutor] felt relevant to me.
Experience of mastery	I felt capable while playing the game [learning from the tutor].
Experience of appropriate challenge	The game [tutor] was challenging but not too challenging.
Situational interest	The game [tutor] was exciting.
Enjoyment	I enjoyed the challenge of learning the material.
Evaluation apprehension	If I did poorly on this activity, people would look down on me.
State anxiety	During this activity, I felt very nervous.

Results

First, we assessed whether students learned from all three conditions. A repeated-measures ANOVA showed a significant improvement for all conditions between pretest and posttest score, $F(1, 338) = 109.35$, $p < .001$, $n_p^2 = .244$. A repeated-measures ANOVA also showed students significantly improved from pretest to delayed posttest score, $F(1, 338) = 191.11$, $p < .001$, $n_p^2 = .361$, confirming that students in all conditions on average performed better on the decimal tests after engaging with the instructional materials.

Next, we examined the correlations between gender dimensions to better understand the relations between binary gender and gender-typed characteristics. Our results showed that female binary gender was positively correlated with feminine-typed characteristics ($r = 0.53$, $p < .001$) and negatively correlated with masculine-typed characteristics ($r = -0.42$, $p < .001$), indicating that students who identified as girls tended to more strongly endorse feminine-typed items and students who identified as boys tended to more strongly endorse masculine-typed items. Feminine-typed characteristics were also positively correlated with masculine-typed characteristics ($r = 0.19$, $p < .001$), demonstrating that many students who endorsed feminine-typed items also tended to endorse masculine typed items, and that such measures are neither exhaustive nor mutually-exclusive. The magnitude of the correlation coefficients indicate that, while these three dimensions of gender were moderately correlated, they were not redundant. Thus, we examined the influence of all three dimensions on learning and enjoyment in our analysis of our research questions.

RQ1: Does the relationship between gender and learning outcomes differ depending on the game narrative?

Table 2 reports the test performance by condition and gender identity. A two-way ANOVA using gender and condition to predict pretest performance showed no main effect of condition at pretest, $F(2, 327) = 1.26$, $p = .29$, $\eta_p^2 = .008$, indicating that students assigned to game and tutor conditions started with the same prior knowledge (i.e. that random assignment to condition was successful). There was a significant main effect of gender identity at pretest, $F(1, 327) = 7.27$, $p = .007$, $\eta_p^2 = .022$, with boys performing better than girls. There was no interaction between gender identity and condition, $F(1, 326) = 0.06$, $p = .81$, $n_p^2 < .001$. A two-way ANCOVA using gender identity and condition to predict posttest performance, controlling for pretest, showed no main effect of condition, $F(2, 326) = 0.81$, $p = .45$, $n_p^2 < .005$, no main effect of gender identity, $F(1, 326) = 2.34$, $p = .13$, $n_p^2 < .007$, and no gender-condition interaction, $F(2, 326) = 0.38$, $p = .69$, $n_p^2 = .002$. Likewise, for delayed posttest performance, a two-way ANCOVA controlling for pretest revealed no main effect of condition, $F(2, 326) = 0.41$, $p = .67$, $n_p^2 = .002$, no main effect of gender identity, $F(1, 326) = 0.010$, $p = .91$, $n_p^2 < .001$, and no gender-condition interaction, $F(2, 326) = 0.57$, $p = .56$, $n_p^2 = .004$. Thus, our predictions that girls would learn more than boys in *Decimal Point* and that boys would learn more than girls in *Ocean Adventure* were not supported; instead, girls and boys learned equally well across both game narratives and in the tutor.

Table 2. Descriptive statistics of test score by gender and condition, reported in M (SD) format.

Test type	*Decimal Point*		*Ocean Adventure*		Tutor	
	Boys (N = 62)	Girls (N = 58)	Boys (N = 43)	Girls (N = 45)	Boys (N = 64)	Girls (N = 61)
Pretest	23.65 (10.17)	21.26 (10.29)	24.86 (12.22)	20.80 (9.22)	25.97 (11.50)	22.98 (9.50)
Posttest	27.00 (9.65)	25.00 (9.08)	29.44 (11.66)	25.24 (8.20)	29.73 (10.08)	26.02 (9.29)
Delayed Posttest	28.27 (9.23)	26.24 (8.91)	30.05 (11.20)	26.49 (9.15)	29.81 (11.05)	28.66 (9.26)

The analyses reported above relied on a measure of self-reported binary gender identity. We also assessed the effects of gender on learning across conditions using gender-typed characteristics. Given that gender-type scales assess characteristics on a spectrum, this approach allowed us to include the six students who identified as non-binary or declined to provide their gender and thus had to be excluded from categorical analyses based on binary gender. We predicted pretest scores with gender-typed characteristics and game condition using multiple regression. Specifically, we built a model that contained masculine-typed characteristics and feminine-typed characteristics, dummy-coded variables for game condition (with the tutor condition as the reference group), and interactions between each gender-typed characteristic dimension and the game conditions. The model was not significant, $F(8, 330) = 1.48$, $p = .16$. The R^2_{adj} value was .011, indicating that 1.1% of the variance in the pretest score was explained by the model. Similarly, to assess the effects of gender-typed characteristics and condition on learning outcomes, we built a multiple regression model with pretest score as covariate and the following predictor variables:

- masculine-typed characteristics,
- feminine-typed characteristics,
- dummy-coded variables for game condition with the tutor as the reference group, and
- interaction terms between each gender-typed characteristic dimension and the condition variables

This model was used to predict posttest and delayed posttest performance. The model predicting posttest scores was significant, $F(9, 329) = 65.03$, $p < .001$. The R^2_{adj} value was .65, indicating that 65% of the variance in posttest score was explained by the model. Within the model, pretest was a significant, positive predictor ($\beta = .79$, $p < .001$), being in the *Decimal Point* condition was a significant, negative predictor ($\beta = -.62$, $p = .022$), and feminine-typed characteristics were a significant, negative predictor ($\beta = -.13$, $p = .014$). No other variables in the model significantly predicted posttest scores. The model predicting delayed posttest scores was also significant, $F(9, 329) = 66.41$, $p < .001$, and the R^2_{adj} value was .64, indicating that 64% of the variance in delayed posttest score was explained by the model. Within the model, pretest was a significant, positive predictor ($\beta = .79$, $p < .001$), and masculine-typed characteristics were a significant, negative predictor ($\beta = -.13$, $p = .034$). Overall, these results differ from the results using the binary measure of gender identity in the emergence of game condition (for posttest) and gender-typed characteristics (for posttest and delayed posttest) as significant predictors. No other predictors in the model were significant.

RQ2: Does the relationship between gender and enjoyment differ depending on game narrative?

We performed a series of two-way ANOVAs assessing the effects of gender identity and condition on each enjoyment construct (affective engagement, experience of meaning, experience of mastery, experience of appropriate challenge, situational interest, and enjoyment; see Table 3 for means and standard deviations). There was a significant main effect of gender identity on affective engagement and experience of mastery, with boys reporting higher levels of affective engagement and experience of mastery than girls. Additionally, we observed a significant main effect of condition for situational interest; post-hoc analyses using the Tukey HSD test indicated that *Decimal Point* led to higher situational interest than the tutor ($T = 2.76$, $p = .017$) and that *Ocean Adventure* led to higher situational interest than the tutor as well ($T = 3.09$, $p = .007$). There was also a significant interaction effect between gender identity and condition on experience of mastery (Figure 6). Additional post-hoc analyses using pairwise T-tests indicated that boys reported significantly higher experience of mastery than girls with *Ocean Adventure* over *Decimal Point* (T = 2.67, p = .009) and with *Ocean Adventure* over the tutor (T = 2.53, p = .013), but not between *Decimal Point* and the tutor (Figure 6). There were no other significant main effects or interaction effects (Table 4).

Table 3. Descriptive statistics of affective experience rating by condition and gender identity, reported in M (SD) format.

Affective Experience	*Decimal Point*		*Ocean Adventure*		Tutor	
	Boys (N = 62)	Girls (N = 58)	Boys (N = 43)	Girls (N = 45)	Boys (N = 64)	Girls (N = 61)
Affective engagement	3.29 (1.14)	2.92 (1.04)	3.49 (1.09)	3.02 (1.02)	3.07 (1.05)	2.89 (0.81)
Experience of meaning	2.96 (1.11)	2.55 (1.14)	2.86 (1.10)	2.80 (0.97)	2.77 (0.87)	2.97 (1.03)
Experience of mastery	3.19 (1.21)	2.23 (1.09)	3.77 (0.98)	3.07 (1.07)	3.25 (1.08)	3.167 (0.93)
Experience of appropriate challenge	3.38 (1.11)	3.30 (0.89)	3.37 (0.97)	3.18 (0.86)	3.19 (0.83)	3.38 (0.83)
Situational interest	3.11 (1.22)	3.08 (1.14)	3.27 (1.15)	3.11 (1.17)	2.49 (1.01)	2.90 (0.96)
Enjoyment	3.29 (1.26)	3.02 (1.18)	3.28 (1.16)	3.12 (1.08)	2.94 (0.99)	3.03 (0.83)
Evaluation apprehension	2.15 (1.01)	2.34 (0.88)	2.08 (0.98)	2.61 (1.05)	2.09 (0.94)	2.49 (1.01)
State anxiety	2.32 (0.96)	2.64 (0.92)	2.39 (0.97)	2.93 (0.99)	2.61 (1.09)	3.01 (0.93)

[Figure 6 about here]

Figure 6. The interaction effect between gender identity and condition on experience of mastery.

Table 4. Results of two-way ANOVAs assessing effects of gender identity and condition on affective experience

Affective Experience	Main Effect of Gender Identity	Main Effect of Condition	Interaction Effect

	F	p	n_p^2	F	p	n_p^2	F	p	n_p^2
Affective Engagement	9.00	.003	.027	1.83	.161	.011	0.51	.602	.003
Situational Interest	0.34	.342	.001	6.07	.003	.037	1.91	.150	.012
Enjoyment	0.87	.351	.003	1.22	.296	.008	0.85	.429	.005
Experience of Meaning	0.58	.449	.002	0.34	.712	.002	2.55	.079	.016
Experience of Mastery	4.24	.040	.013	1.24	.291	.008	3.25	.040	.020
Experience of Appropriate Challenge	0.05	.817	<.001	1.33	.267	.008	0.08	.927	<.001
Evaluation Apprehension	11.21	.001	.035	0.40	.673	.003	0.87	.422	.006
State Anxiety	14.16	<.001	.043	3.36	.036	.021	0.30	.739	.002

The analyses reported above examined the effects of game condition and gender using self-reported gender identity. To investigate the same research question using gender-typed characteristics, we built multiple regression models predicting each of the constructs based on masculine-typed characteristics, feminine-typed characteristics, dummy-coded variables for game condition (with the tutor condition as the reference group), and interactions between each gender-typed characteristic dimension and the game conditions. Of the models, all models except those for affective engagement and experience of appropriate challenge were statistically significant (Table 5). Masculine-typed characteristics were a significant, positive predictor within the models for situational interest ($\beta = 0.51$, $p = .017$), enjoyment ($\beta = 0.67$, $p = .002$), experience of meaning ($\beta = 0.59$, $p = .005$), and experience of mastery ($\beta = 0.43$, $p = .042$). Additionally, feminine-typed characteristics were a significant, positive predictor within the models for situational interest ($\beta = 0.51$, $p = .044$) and experience of mastery ($\beta = 0.60$, $p = .015$). Gender-typed characteristics were not significant in any other models; condition and the interactions between condition and gender-typed characteristics were also not significant in any models.

Table 5. Statistical Significance of Multiple Regressions

Linear Regression Model	R^2_{adj}	F	p
Affective Engagement	.020	$F(8,330) = 1.84$	.069
Situational Interest	.108	$F(8,330) = 5.90$	<.001
Enjoyment	.057	$F(8,330) = 5.50$	<.001
Experience of Meaning	.037	$F(8,330) = 2.60$	.009
Experience of Mastery	.077	$F(8,330) = 4.45$	<.001
Experience of Appropriate Challenge	.020	$F(8,330) = 1.82$	.073
Evaluation Apprehension	.051	$F(8,330) = 3.19$	.002

State Anxiety	.052	$F(8,330) = 3.22$	.002

RQ3: Do girls and students with more feminine-typed characteristics experience lower levels of anxiety when learning math through digital learning games compared to tutor instruction?

To test this research question, we focused on two anxiety-related constructs: state anxiety and evaluation apprehension. We conducted a two-way ANOVA assessing each construct as an outcome with condition and binary gender identity as predictors (Table 4). For evaluation apprehension, there was a main effect of gender identity, with girls reporting higher levels of evaluation apprehension than boys. There was no main effect of condition, and no interaction. Finally, a two-way ANOVA predicting state anxiety revealed a significant main effect of gender identity, with girls reporting higher levels of anxiety than boys. There was also a significant effect of game condition, and post-hoc analysis using a Tukey HSD test found that *Decimal Point* led to lower levels of state anxiety than the tutor ($T = -2.59$, $p = .027$). There was no significant interaction between gender and condition.

Finally, we also tested this research question using gender-typed characteristics. We built multiple regression models predicting evaluation apprehension and state anxiety with masculine-typed characteristics, feminine-typed characteristics, dummy-coded variables for game condition (with the tutor condition as the reference group), and interactions between each gender-typed characteristic dimension and the game conditions. Both models were statistically significant (Table 5). In both models, the interaction term for feminine-typed characteristics and the tutor condition were significant, indicating that in the tutor condition, feminine-typed characteristics were a significant, positive predictor of evaluation apprehension ($\beta = 0.77$, $p = .018$) and state anxiety ($\beta = 0.66$, $p = .043$). No other predictors in either model were significant.

Discussion and Conclusion

In this study, we investigated potential gender effects across three learning platforms with identical instructional materials: a digital learning game that has yielded consistent learning advantages for girls over boys, a new game with a more masculine-oriented narrative, and a conventional tutor. Our research investigated two potential pathways through which games might influence learning. First, we investigated whether math games reduced the negative affective experiences (e.g., math anxiety, evaluation apprehension) that girls tend to experience in math. Second, we investigated the degree to which games promote positive affective experiences (e.g., affective engagement, enjoyment, situational interest) and, specifically, whether students benefit more when the game narrative aligns with gendered narrative preferences. In this setting, we also incorporated a multidimensional model of gender to examine its advantages over the traditional binary model of gender identity in predicting students' learning and playing experience. Overall, our findings yield insights into how games influence students' affective experiences across different dimensions of gender, which have important implications for the design of learning games that help bridge gender gaps in STEM education and participation. We further discuss these insights as follows.

First, our findings with respect to gender differences in learning are not consistent with those reported in prior studies of *Decimal Point* and the decimal tutor. In previous studies, girls had lower pretest scores than boys and caught up at posttest and delayed posttest (Nguyen et al., 2022). Instead, our results indicate that all conditions were equally beneficial for short- to mid-term learning gain. Finer grained analyses of students' learning over time, for example with knowledge tracing and learning curves (Stamper & Koedinger, 2011), could identify specific learning activities that induced the gender effects found in prior studies or inhibited it in this study. Such knowledge can inform instructional and game redesign to ensure learning benefits for all students.

In terms of students' affective experience, we found that girls reported greater evaluation apprehension and state anxiety than boys, which is consistent with many previous studies showing that salient math content can induce the experience of anxiety, especially among students with stronger feminine gender identities (Schmader, 2002). Importantly, feminine-typed characteristics were a significant, positive predictor of anxiety and evaluation apprehension in the tutor condition only. In other words, both games appeared to eliminate the tendency for students with stronger feminine identities to experience math anxiety. It may be that the playful game environments reduce the saliency of the math work students are doing, which in turn could reduce anxiety for students whose gender identity makes them sensitive to cues of math saliency. The fact that this difference emerged only with gender-typed characteristics and not binary gender is consistent with some prior research showing that gender stereotypes around math have a stronger effect on girls with stronger feminine identities (Schmader, 2002). Given that the interaction between gender and learning platform on anxiety and evaluation apprehension was uncovered with analyses of gender-typed characteristics but not gender identity, we recommend incorporating the multidimensional gender framework in future research to reveal subtle and nuanced patterns of gender differences in student affect that crude measures of binary gender might not capture.

Across the three learning platforms, we found that boys reported greater affective engagement. Interestingly, we also identified a significant effect of the masculine game narrative, which led boys to report greater experience of mastery than girls in the *Ocean Adventure*. Higher experience of mastery indicated that boys felt more capable during game play, which likely results from a stronger sense of accomplishment offered by the new narrative (Eseryel et al., 2014), where the player went from solving decimal exercises to defeating pirate leaders and eventually saving the world. Thus, our results suggest that switching to a masculine game narrative can promote boys' engagement, which may contribute to the development of interest (Harackiewicz et al., 2016). At the same time,

Ocean Adventure did not affect boys and girls differently on any other measures of positive affective experiences, although there was an overall effect of both games promoting greater situational interest than the tutor. To our knowledge, this is the first research which examines two separate narratives for the same learning game. While overall narrative did not appear to have strong effects on learners' experiences or outcomes, we did find some indication of boys engaging differently with a more masculine narrative. Future research should evaluate how narrative manipulation in other learning games may affect students' learning experiences based on gender as well as other individual differences that may more strongly relate to narrative (e.g., narrative preferences).
Our analyses of multiple dimensions of gender (i.e., gender identity as well as masculine-typed and feminine-typed characteristics) further revealed several nuances to the gender differences in learning and enjoyment. First, gender identity was a significant predictor of pretest scores, but not a significant predictor of posttest and delayed posttest scores. Meanwhile, the gender-typed characteristics were significant predictors of posttest and delayed posttest scores. Interestingly, both masculine-typed and feminine-typed characteristics were negative predictors of performance. This is consistent with the prediction that the more nuanced, continuous measures of gender typing might be able to detect variability that is lost with the dichotomous measure of gender identity. Alternatively, it could suggest that gender-typed characteristics, as measured here with Liben and Bigler's (2002) scale, are more relevant to learning processes and outcomes than binary gender identity. In addition, we found that masculine-typed characteristics were a significant and positive predictor of situational interest, enjoyment, and experience of mastery. At the same time, feminine-typed characteristics were also a significant and positive predictor of situational interest and experience of mastery.

While a prior study of *Decimal Point* from 2017 has shown that it leads to more learning and enjoyment than the tutor (McLaren, Farzan, et al., 2017), the game's learning advantage was not replicated in our current study. Given a fairly large time gap between the two studies, with the COVID-19 pandemic in between that had significant impacts on students' math performance (Contini et al., 2022) and game play behavior (Kourti et al., 2021), the lack of replication may stem from a variety of pedagogical and motivational factors that should be investigated in future research. Nonetheless, we found that the games were more engaging to students than the tutor, as indicated by their reports of situational interest, which suggests there are still measurable benefits to incorporating digital learning games in the classroom.

Finally, we acknowledge several limitations of the current work. First, while students were randomly assigned to either *Decimal Point* or *Ocean Adventure*, we used classroom-level assignment to games vs. tutor conditions. We adopted this approach to avoid the potential for distraction or learning disruption if some students were playing a game and others were using a tutor in the same classroom. However, schools may have unevenly distributed students according to prior knowledge; although all schools stated that they did not use tracking (i.e., grouping students in different classes based on performance or perceived ability), some teachers anecdotally mentioned knowledge differences across their classes. Such uneven distribution of students' prior knowledge or educational needs may have affected results. However, there were no effects of condition on pretest performance, suggesting that students across conditions were nevertheless equivalent in their prior knowledge. Finally, assignment between the games was skewed towards *Decimal Point*, with 147 students assigned to play *Decimal Point* and only 117 assigned to play *Ocean Adventure*. Although sufficient samples were acquired for each condition, the imbalance of assignment to conditions may have interfered with our ability to make observations particularly regarding *Ocean Adventure*. Careful attention to properly distributing students across conditions will be necessary in subsequent studies. Finally, we did not assess the effects of the classroom environment in our current analyses.

In conclusion, this research reveals important insights into the role of gender and game environment in shaping students' experience with learning games. We found that digital learning games may be especially effective at reducing anxiety and evaluation apprehension for students with feminine-typed characteristics in domains with gender stereotypes such as math. We also found that adapting games to gendered preferences can help promote some aspects of engagement, but this practice should be based on empirical research rather than intuition. In addition, utilizing a multidimensional model of gender results in a more inclusive and nuanced understanding of the role of gender in game-based learning. These findings yield important lessons into the design of inclusive and effective digital learning games and provide important features for modeling students' learning and enjoyment to contribute to our understanding of design considerations for educational games.

Acknowledgements

This work was supported by the National Science Foundation Award #DRL2201796. The opinions expressed are those of the authors and do not represent the views of NSF. Thanks to Jimit Bhalani, John Choi, Kevin Dhou, Darlan Santana Farias, Rosta Farzan, Jodi Forlizzi, Craig Ganoe, Rick Henkel, Scott Herbst, Grace Kihumba, Kim Lister, Patrick Bruce Gonçalves McLaren, and Jon Star for important contributions to the development and early experimentation with *Decimal Point*. Thanks also to Rick Henkel, Sarah Di, Em Tyminski, Kevin He, and Leah Teffera for important contributions to the development and early experimentation with *Ocean Adventure*.

References

Arroyo, I., Burleson, W., Tai, M., Muldner, K., & Woolf, B. P. (2013). Gender differences in the use and benefit of advanced learning technologies for mathematics. *Journal of Educational*

Psychology, *105*(4), 957–969. https://doi.org/10.1037/a0032748

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, *147*(2), 134–168. https://doi.org/10.1037/bul0000307

Ben-Eliyahu, A., Moore, D., Dorph, R., & Schunn, C. D. (2018). Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement across science activities and contexts. *Contemporary Educational Psychology*, *53*, 87–105.

Breda, T., Jouini, E., & Napp, C. (2018). Societal inequalities amplify gender gaps in math. *Science*, *359*(6381), 1219–1220. https://doi.org/10.1126/science.aar2307

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243.

Chou, C., & Tsai, M.-J. (2007). Gender differences in Taiwan high school students' computer game playing. *Computers in Human Behavior*, *23*(1), 812–824.

Chung, L.-Y., & Chang, R.-C. (2017). The effect of gender on motivation and student achievement in digital game-based learning: A case study of a contented-based classroom. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(6), 2309–2327.

Contini, D., Di Tommaso, M. L., Muratori, C., Piazzalunga, D., & Schiavon, L. (2022). Who lost the most? Mathematics achievement during the COVID-19 pandemic. *The BE Journal of Economic Analysis & Policy*, *22*(2), 399–408.

Cook, R. E., Nielson, M. G., Martin, C. L., & DeLay, D. (2019). Early adolescent gender development: The differential effects of felt pressure from parents, peers, and the self. *Journal of Youth and Adolescence*, *48*, 1912–1923.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103.

Else-Quest, N. M., Mineo, C. C., & Higgins, A. (2013). Math and Science Attitudes and Achievement at the Intersection of Gender and Ethnicity. *Psychology of Women Quarterly*, *37*(3), 293–309. https://doi.org/10.1177/0361684313480694

Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An investigation of the interrelationships between motivation, engagement, and complex problem solving in game-based learning. *Journal of Educational Technology & Society*, *17*(1), 42–53.

Farrell, D., & Moffat, D. C. (2014). Adapting cognitive walkthrough to support game-based learning design. *International Journal of Game-Based Learning (IJGBL)*, *4*(3), 23–34.

Forlizzi, J., McLaren, B. M., Ganoe, C., McLaren, P. B., Kihumba, G., & Lister, K. (2014). *Decimal Point: Designing and Developing an Educational Game to Teach Decimals to Middle School Students*. https://www.semanticscholar.org/paper/Decimal-Point-%3A-Designing-and-Developing-an-Game-to-Forlizzi-McLaren/8c3749bb2294b33339d7396245b2b38845fc6c0f

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (Rev. and updated ed). Palgrave Macmillan.

Ghasemi, E., & Burley, H. (2019). Gender, affect, and math: A cross-national meta-analysis of Trends in International Mathematics and Science Study 2015 outcomes. *Large-Scale Assessments in Education*, *7*(1), 10. https://doi.org/10.1186/s40536-019-0078-1

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest Matters: The Importance of Promoting Interest in Education. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 220–227. https://doi.org/10.1177/2372732216655542

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, *74*(2), 171.

Isotani, S., McLaren, B. M., & Altman, M. (2010). *Towards intelligent tutoring with erroneous examples: A taxonomy of decimal misconceptions*. 346–348.

Jiang, S., Simpkins, S. D., & Eccles, J. S. (2020). Individuals’ math and science motivation and their subsequent STEM choices and achievement in high school and college: A longitudinal study of gender and college generation status differences. *Developmental Psychology*, *56*(11), 2137–2151. https://doi.org/10.1037/dev0001110

Joiner, R., Iacovides, J., Owen, M., Gavin, C., Clibbery, S., Darling, J., & Drew, B. (2011). Digital Games, Gender and Learning in Engineering: Do Females Benefit as Much as Males? *Journal of Science Education and Technology*, *20*(2), 178–185. https://doi.org/10.1007/s10956-010-9244-5

Khan, A., Ahmad, F. H., & Malik, M. M. (2017). Use of digital game-based learning and gamification in secondary school science: The effect on student engagement, learning and gender difference. *Education and Information Technologies*, *22*(6), 2767–2804. https://doi.org/10.1007/s10639-017-9622-1

Klisch, Y., Miller, L. M., Wang, S., & Epstein, J. (2012). The Impact of a Science Education Game on Students' Learning and Perception of Inhalants as Body Pollutants. *Journal of Science Education and Technology*, *21*(2), 295–303. https://doi.org/10.1007/s10956-011-9319-y

Kourti, A., Stavridou, A., Panagouli, E., Psaltopoulou, T., Tsolia, M., Sergentanis, T. N., & Tsitsika, A. (2021). Play behaviors in children during the COVID-19 pandemic: A review of the literature. *Children*, *8*(8), 706.

Levine, S. C., & Pantoja, N. (2021). Development of children's math attitudes: Gender differences, key socializers, and intervention approaches. *Developmental Review*, *62*, 100997. https://doi.org/10.1016/j.dr.2021.100997

Liben, L. S., Bigler, R. S., Ruble, D. N., Martin, C. L., & Powlishta, K. K. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*.

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational* *and Psychological Measurement*, *70*(4), 647–671.

Mayer, R. E. (2014). *Computer Games for Learning: An Evidence-Based Approach*. https://doi.org/10.7551/mitpress/9427.001.0001

Mayer, R. E. (2019). Computer Games in Education. *Annual Review of Psychology*, *70*(Volume 70, 2019), 531–549. https://doi.org/10.1146/annurev-psych-010418-102744

McLaren, B. M. (2024). A Decade of Learning Science Findings with a Digital Learning Game. *In P. Ilic, R. Wegerif, & I. Casebourne (Eds.), Artificial Intelligence in Education: The Intersection of Technology and Pedagogy. Springer Verlag Germany*.

McLaren, B. M., Adams, D. M., Mayer, R. E., & Forlizzi, J. (2017). A Computer-Based Game that Promotes Mathematics Learning More than a Conventional Approach: *International Journal of Game-Based Learning*, *7*(1), 36–56. https://doi.org/10.4018/IJGBL.2017010103

McLaren, B. M., Farzan, R., Adams, D., Mayer, R., & Forlizzi, J. (2017). *Uncovering gender and problem difficulty effects in learning with an educational game*. 540–543.

McLaren, B. M., Richey, J. E., Nguyen, H. A., & Mogessie, M. (2022). A digital learning game for mathematics that leads to better learning outcomes for female students: Further evidence. *ECGBL 2022 16th European Conference on Game-Based Learning*. https://books.google.com/books?hl=en&lr=&id=INOYEAAAQBAJ&oi=fnd&pg=PA339&dq=McLaren,+B.M.,+Richey,+J.E.,+Nguyen,+H.A.,+%26+Mogessie,+M.+(2022).+A+digital+learning+game+for+mathematics+that+leads+to+better+learning+outcomes+for+female+students:+Further+evidence.+In:+Proceedings+of+the+16th+European+Conference+on+Game+Based+Learning+(ECGBL+2022).&ots=Ge0U0-cmDS&sig=X1jlxc-nZbtcTrXO7fRHF2-3v9k

Mejía-Rodríguez, A. M., Luyten, H., & Meelissen, M. R. M. (2021). Gender Differences in Mathematics Self-concept Across the World: An Exploration of Student and Parent Data of TIMSS 2015. *International Journal of Science and Mathematics Education*, *19*(6), 1229–1250. https://doi.org/10.1007/s10763-020-10100-x

Nguyen, H. A., Else-Quest, N., Richey, J. E., Hammer, J., Di, S., & McLaren, B. M. (2023). *Gender Differences in Learning Game Preferences: Results Using a Multi-dimensional Gender Framework*. 553–564.

Nguyen, H. A., Hou, X., Richey, J. E., & McLaren, B. M. (2022). The Impact of Gender in Learning With Games: A Consistent Effect in a Math Learning Game. *International Journal of Game-Based Learning*, *12*(1), 1–29. https://doi.org/10.4018/IJGBL.309128

Niederle, M., & Vesterlund, L. (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of Economic Perspectives*, *24*(2), 129–144. https://doi.org/10.1257/jep.24.2.129

Pekrun, R. (2005). Progress and open problems in educational emotion research. *Learning and Instruction*, *15*(5), 497–506.

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating Gender Differences in Mathematics and Science: Results from the 2011 Trends in Mathematics and Science Survey. *Research in Science Education*, *49*(1), 25–50. https://doi.org/10.1007/s11165-017-9630-6

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, *38*(2), 194–201. https://doi.org/10.1006/jesp.2001.1500

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4–28.

Stamper, J. C., & Koedinger, K. R. (2011). *Human-machine student model discovery and improvement using DataShop*. 353–360.

Tsai, F.-H. (2017). An Investigation of Gender Differences in a Game-based Learning Environment with Different Game Modes. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(7), 3209–3226. https://doi.org/10.12973/eurasia.2017.00713a

Van Mier, H. I., Schleepen, T. M. J., & Van den Berg, F. C. G. (2019). Gender differences regarding the impact of math anxiety on arithmetic performance in second and fourth graders. *Frontiers in Psychology*, *9*. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02690

Veit, C. T., & Ware, J. E. (1983). Mental health inventory. *Psychological Assessment*.

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30-year examination. *Intelligence*, *38*(4), 412–423. https://doi.org/10.1016/j.intell.2010.04.006