Article

Not peer-reviewed version

A Math Digital Game and Multiple-Try Use in Primary Students: A Gender Analysis on Motivation and Performance

Claudio Cubillos *, Silvana Roncagliolo, Daniel Cabrera-Paniagua, Rosa Maria Vicari

Posted Date: 18 April 2024

doi: 10.20944/preprints202404.1186.v1

Keywords: Computer-Based Learning; Sex differences; Multiple attempts; MTF

Article

A Math Digital Game and Multiple-Try Use in Primary Students: A Gender Analysis on Motivation and Performance

Claudio Cubillos [1,*], Silvana Roncagliolo [1], Daniel Cabrera-Paniagua [2] and Rosa Maria Vicari [3]

[1] Escuela de Ingeniería Informática, Pontificia Universidad Católica de Valparaíso, Valparaíso 2362807, Chile; silvana.roncagliolo@pucv.cl (S.R.)
[2] Escuela de Ingeniería Informática, Universidad de Valparaíso, Valparaíso 2362905, Chile; daniel.cabrera@uv.cl (D.C.-P.)
[3] Federal University of Rio Grande do Sul, Porto Alegre, Brazil; rosa@inf.ufrgs.br (R.M.V.)
* Correspondence: claudio.cubillos@pucv.cl (C.C.)

Abstract: Gender has been a rarely addressed aspect in digital game-based learning (DGBL). Likewise, mixed results have been presented regarding the effects according to gender and the conditions that generate these effects. The present work studies the effects of a drill-and-practice mathe-matical game on primary students. The study focused on an analysis by gender, measuring learning performance, and motivation in the practice activity. Also, two instructional mechanics were considered regarding the question answering, a multiple-try (MTF) and a single-try (STF) condition, to search for possible differences. A total of 81 students from 4 courses and 2 schools participated in the experiment. The study's main findings were that: (a) the girls outperformed boys in terms of students' learning gains, (b) girls presented lower levels of competence and autonomy than boys, (c) in the MTF, girls presented lower levels of autonomy but no differences in competence contrasted to boys, (d) in the STF, girls presented lower levels of competence but no differences in autonomy against boys, and (e) no differences existed in interest, effort and value among gender overall o per instructional condition. This study enhances the knowledge of gender differences under diverse instructional settings, in particular providing insights into possible differences by gender when varying the number of attempts provided to students.

Keywords: computer-based learning; sex differences; multiple attempts; MTF

1. Introduction

In recent decades, rapid technological advancements, particularly in computing and information technology (IT), have invaded nearly every facet of daily life, including education. Nowadays various IT tools exist to support or complement learning processes, including Learning Management Systems (LMS) like Chamilo, Moodle, and H5P, platforms for building and offering courses (e.g., Udemy), educational game generators (e.g., Educandy, Educaplay), and educational math platforms like photoMath or Symbolab. Overall, the effect of technology on learner motivation, attitudes, and achievement showed to be significant, especially when learning mathematics [1].

Gender differences exist in diverse aspects related to technology use and learning. Males are more likely to be interested in using computers and technologies than females [2,3]. They also perceive it as more beneficial to use multimedia technology for learning than females [3,4]. Furthermore, while males tend to focus on usefulness, females pay more attention to enjoyment and ease of use [4–6]. For example, a recent study on game use of Pokémon-Go [5] found that females tended to play for enjoyment, whereas males were looking for achievement and social interactivity.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

On cognition and learning, studies suggest that girls perform better than boys in planning and attention tasks [7,8]. Additionally, it seems that men and women employ distinct strategies in planning tasks, with women having higher verbal and memory skills, while men exhibit stronger capabilities in spatial cognition and learning, and both develop different approaches to complex planning tasks throughout their lives [8,9]. Therefore, gender is a key factor if we expect to use educational digital technologies efficiently and effectively.

Digital Game-Based Learning has extensively researched various emotional and motivational factors due to their significant relevance in the learning process. A positive relationship exists between engagement and performance [10–12], while engagement in games predicts in-game students' performance and impacts learning outcomes [13]. Thus, including motivational and emotional aspects in research studies is necessary to have a clearer picture of the environmental factors and students' internal processes that affect learning.

Studies exist covering gender differences in learning gains and engagement [7,14], attitudes [15], motivation [16], and anxiety [17,18]. Some show contradictory findings, and some others do not provide fine-grain relations among gender, motivation, and other emotional variables influencing learning. Hence, diverse past studies agree on the need for further research that incorporates gender, motivation, and learning [14,19].

Integrating gaming with learning is not an easy task, as even setting up any of the above-mentioned learning platforms is not straightforward. Many of these tools allow for the configuration of different instructional features such as question types, number of attempts, and types of feedback. However, a key problem in their use is that it is not always clear which combinations of all these available characteristics are most suitable for a specific group of students.

Adding the factor of gender differences makes it even less clear which of these characteristics generate gender-neutral effects and which do not. Therefore, there is a lack of gender studies that include instructional factors such as the item and feedback type, together with the number of attempts.

Based on the above, the present study aims to explore the performance and motivational effects of using a math digital game on primary students, focusing on searching for possible differences by gender, especially when varying the number of attempts provided as part of the instructional design.

In the following, the next section provides a background on existing literature covering gender in learning digital contexts, and the inclusion of motivation and other emotional outcomes, for then deepening the concept of instructional feedback and multiple attempts. Then, a sub-section on self-determination theory is depicted to frame the motivational constructs to use in our study. Section 3 presents the objectives and research questions guiding the current study together with a detailed description of the experiment's materials and methods, including the MatematicaST game, the participants, instruments, and the procedure used. Then, Sections 4 and 5 show the results and discussion, for then end up concluding in section 6 with the main findings and their implications.

2. Background

2.1. Gender and Learning in Digital Contexts

The study focuses on the biological sex of elementary school students rather than their sexual identification or orientation. Therefore, the terms gender and sex are used interchangeably.

Males and females approach technology in different ways. Teen girls spend more time on smartphones and being online on social networks, while boys spend more time on electronic devices playing games [20]. However, in a mobile learning setting, males are more influenced by social factors than females [21].

Boys consider games unique, engaging experiences, while females consider them not as memorable or entertaining, but just as an alternative way to learn [22]. Males perceive games as more relaxing than females and that allows them to work better. The association of diversion and catharsis to games is higher in males. Females are always more skeptical than males about the instructional potential of games [22].

 doi:10.20944/preprints202404.1186.v1

Regarding gender differences in students' academic performance in schools, studies show that females tend to have better academic achievements [23] and outperform males on learning assignments [24]. On students' disposition towards risk, Cipriani [25] examined the effects of multiple attempts in multiple-choice tests. The study reported that females tended to omit more items and guess less than males because men are more risk-prone and interpret risk as a challenge rather than a threat.

Despite these differences, few studies have investigated gender differences in gameplay and DGBL [14]. Some research [14,26,27] suggests that females perform better than males under DGBL in terms of engagement and learning gains. For example, in Lukosch et al. [7] female participants outperformed male participants in a game to develop planning operations skills in a container terminal setting. Yeo et al. [31] examined the effects of gender and prior knowledge on students' learning performance and motivation using a food-chain digital game on 5th-grade students in Taiwan, showing that improvements in learning performance were higher in females than males. The study by Khan et al. [14] considered a game-based learning (GBL) chemistry (reactivity) application for 8th-grade Pakistani students. They measured the impacts on learning, engagement, and differences among genders. They found that girls outperformed males in post-test scores under the educational application while no gender differences were found under the traditional science instruction group. In addition, no differences among control and treatment groups were found in the pre- and post-test scores, providing more support to the idea of learning differences due to gender.

Meanwhile, other research [16,19,28] reported no significant learning differences by gender. The study of Chung & Chang [16] involved a digital game for first aid training (EFA) in English for Chinese non-native speakers, following a content-based instruction (CBI) approach. Their findings showed no significant differences existing by gender in learning achievements. Subsequently, it is necessary to clarify which conditions favor learning in one gender over the other and which do not.

2.2. Gender, Motivation, and Emotional Outcomes

Regarding gender and student attitudes, in Bahar & Asil [15], males reported more positive attitudes than females towards e-assessments. On students' interest in educational video games, Manero, Torrente, Fernández-Vara, & Fernández-Manjón [29] explored the impact of gender, age, and gaming habits on the effectiveness of the game "La Dama Boba" (The Foolish Lady), a graphical adventure based on the theater play of the same name. It showed no gender differences in students' interest in theater-going.

However, in the study of the digital game for first aid training (EFA) by Chung & Chang [16], girls presented higher motivation levels than boys, although similar learning levels. Authors argue that their game was moderate genre as the three learning activities used elements such as storylines, challenges that did not emphasize competition, educational values, fun factors, and interactivity that offered game exploration opportunities, enhancing female motivation. They also measured usability with a SUS test, finding a non-significant higher score on females. The study did not compare or analyze gender effects per group or condition to assess possible gender differences with and without game conditions.

Khan et al. [14] study measured students' engagement with 4 factors: positive body language, consistent focus, confidence, and fun and excitement. They assessed significant differences in these factors among control and treatment conditions. Girls in the learning application presented higher engagement levels than boys, although not significant. Furthermore, the engagement instrument was the "student engagement walk-through checklist" [30] and was administered by teachers during the intervention.

In the experience with the food-chain digital game of Yeo et al. [31] with primary students, they used the ARCS Learning Motivation Scale [32] which is based on the expectation value theory and the individuals' success expectations. It considers 4 constructs: attention, relevance, confidence, and satisfaction. They found that the attention of the medium and high prior knowledge (PK) groups was significantly higher than the low PK group. Also, the medium PK group reported higher satisfaction levels than the low and high PK groups. Results also varied by gender. In the high PK group males

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

presented higher attention levels than females, while in the low PK group, males showed higher confidence than females.

Besides, a recently explored variable has been student anxiety, especially in math [33,34] and in digital-based game learning [35]. Diverse authors [17,36] found no gender differences concerning anxiety and achievement in math.

In Goetz et al. [37] examined possible gender differences between trait (habitual) anxiety and state (transitory) anxiety in mathematics in nearly 700 students. They concluded that females presented a higher trait of math anxiety compared to males, but no differences existed for state anxiety using experience-sampling methods in activities like taking a math test or attending math lessons. Another relevant finding was that students' self-perceptions of their competence in mathematics were lower in girls than boys despite having similar grades in math. Authors argue that this difference in perceived competence helps to partially explain gender differences among state-trait anxiety. More recently, Wang [18] suggests that the student's spatial ability can be a relevant factor in mediating math anxiety when considering gender differences.

Therefore, the effects of gender on learning performance, engagement, motivation, and anxiety are not straightforward. Different variables and conditions make it hard to draw consistent conclusions from the revisited literature. Diverse authors [14,19,26] highlight that more studies are needed focusing on gender, learning approaches, engagement, and achievement for building strategies that foster gender parity in DGBL.

2.3. Instructional Feedback and Multiple Attempts

Instructional feedback can be defined as providing students with information to correct their answers, aiding in error identification, misconception correction, and the enhancement of problem-solving strategies and self-regulation [38–40]. In a broader sense, feedback encompasses any information provided by an agent—such as a teacher, peer, book, parent, or experience—regarding a person's performance or understanding [41]. In this study, instructional feedback is centered on the task level and its formative dimension (learning), diverging from delayed summary feedback [42] and feedback focused on task motivation or self-regulation processes [40].

Feedback provision can vary widely [43], with three feedback types being most frequently used in game-based learning settings during the last decade [40,44,45] as part of single-attempt instructional feedback implementations (STF). Knowledge of Result (KR) confirms the correctness of students' answers or marks errors without additional information [40]. Knowledge of Correct Response (KCR) specifies the correct answer without further explanation [46]. Elaborated feedback (EF) provides information on why an answer is right or wrong, offering explanations, additional materials, hints, or a combination of these [45,46]. Additionally, feedback types differ based on the number of attempts allowed. Answer-Until-Correct (AUC) feedback permits multiple attempts until the correct answer is reached, offering KR feedback between attempts [47]. Multiple-try feedback (MTF) allows a limited number of attempts, typically providing KR but sometimes offering KCR after the final try or hints on the first attempt [48].

Past literature has progressed in leveraging under what conditions each type of feedback seems better, but mainly focused on single-try (STF) alternatives. A study by Van der Kleij, Feskens, & Eggen [39] revealed that much research exists on elaborated feedback combined with KR or KCR under computer-based environments. Their meta-analysis showed that different feedback types were moderated by the kind of learning to be achieved, with EF outperforming both KCR and KR under high–order learning outcomes (LOs) while no major differences existed among these in low-order learning that involved verbatim or recall of information.

The above highlights the relevance of task complexity. By considering that feedback's primary importance is the correction of errors, one would expect to see larger effects for instruction with higher error rates [49], that is, more difficult topics have more possibilities for learning improvement. Differences in single-try feedback alternatives (KR, KCR, and EF) appear only upon student error.

Now regarding multiple-attempts use, there is much less literature and outdated, besides providing mixed results on their learning effects. Some studies show significant learning gains with

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

multiple-try solutions [42,48,50], while others find no significant differences when compared to single-try alternatives [51–53]. Only the review by Clariana & Koul [50] compared MTF to different feedback types (KR, KCR, and EF), obtaining that MTF outperformed the other feedback types for higher-order LOs (Effect Size - ES .11) while being equivalent to or inferior for lower-order outcomes (ES - .22). It seems that in situations where test items involve comprehension and understanding rather than simply recall, such as in mathematical problem-solving, the invitation to try again offers a chance for elaboration and reorganization of information, potentially enhancing learning [48].

A relevant aspect is that the use of multiple trials (AUC and MTF) offers an interactive mechanism [44] for students' errors, giving them multiple immediate exposures to the same item [54]. From a contiguity theory perspective, multiple trials seem beneficial compared to single attempts because may engage learners in additional active processing following errors [55,56]. It is also in line with information processing theory, arguing that the continued engagement with a question needed upon incorrect response can offer potential advantages [57,58] and the idea that providing the correct answer after only one response, as happens with KCR, may "short circuit" learning [55,59].

On the other hand, some authors argue that repeatedly asking a learner to answer the same question can be frustrating [55,60]. Providing multiple attempts at errors might encourage deeper thinking about the lesson unless the learner falls on random guessing because of frustration or impatience [61]. Furthermore, diverse past studies support the idea that learners, particularly those with low academic performance, can feel frustrated when lessons employ multiple-try feedback [61–63].

Referring to Salomon & Globerson's concept of mindfulness[64], feedback can aid learning when received mindfully but might hinder it if it promotes mindlessness[48]. In this line, besides knowing that MTF seems to be more beneficial for high-level learning outcomes, there is not much clarity regarding what other factors and their interactions promote mindful or mindless trial-and-error behavior when using multiple attempts.

2.4. Motivation and Self-Determination

The Self-Determination Theory (SDT) [65,66] is a psychological framework for explaining the motivation construct and applied in DGBL. SDT coins the concept of intrinsic motivation, the innate psychological need for competence, autonomy, and relatedness [67]. Competence involves the need for effectiveness in interactions with the environment [67], while autonomy regards experiencing free choice and "freedom in initiating one's behavior" [68]. Relatedness refers to the fundamental human need for belongingness and connection with others [68]. In addition, the Cognitive Evaluation Theory (CET) (an SDT sub-theory) explores factors that either enhance or hinder intrinsic motivation [66]. CET states that fostering intrinsic motivation in an educational setting involves providing diverse stimulating instructional elements, adequate challenges, and promoting learner initiative and autonomy without control and pressure [65]. In this sense, activities with optimal challenge and effort levels enhance motivation [67]. Adequate challenges lead to effort, generating feelings of competence while the absence of controlling conditions can positively impact effort levels[65]. Therefore, SDT-CET theory makes use of the following constructs to assess motivation: interest, competence, pressure, effort, choice, and value. All of them are positively correlated to motivation except for pressure, the reason why it is often used in reverse mode as no pressure.

Research on game-based learning has used SDT with diverse focuses. Liao, Chen, & Shih [69] studied the relationship between cognitive load, motivation, and learning outcomes. While [70] studied the properties of scaffolds on intrinsic motivation. SDT was also used to explain the relation of gamification techniques with motivation in virtual reality [71] and for bridging learning mechanics and game mechanics under GBL contexts [72]. Liu, Wang, & Lee [73] found that game quality and feedback significantly influenced the motivation and learning performance of students.

In our previous study utilizing MatematicaST [74], the focus was on contrasting multiple-try feedback and a single-try condition with KCR among primary school students. The results revealed that MTF provided higher learning while exhibiting higher levels of perceived competence and autonomy than single try with KCR. Also, multiple-try feedback presented an increase in perceived

 doi:10.20944/preprints202404.1186.v1

pressure. No significant differences were observed in terms of perceived effort and value between conditions and both remained consistently high. As well, no gender analysis was conducted as it wasn't a factor or focus of the paper.

In summary, SDT has been applied to gamification, game mechanics, and game quality in computer-based learning contexts, but without controlling instructional feedback such as single-try and multiple attempts. There is little empirical research examining the SDT constructs with gender and feedback types under DGBL contexts, showing necessary to advance the literature on how gender affects students' engagement, motivation, and other emotional dimensions such as effort and pressure. And also on how feedback type and multiple attempts' inclusion could affect gender outcomes in terms of learning and motivation.

3. Materials and Methods

3.1. The Present Study Objectives and Research Questions

Based on the above, the present study first aims to explore the impacts of gender when using a math digital game on primary students. Second, intends to assess if such impacts differ across two instrucional feedback conditions (namely MTF and STF). The study focuses not only on performance effects (learning gains), but also on motivational effects based on the SDT-CET theory and operationalized in terms of interest, competence, autonomy, effort, no pressure, and value constructs. Three high-order math learning objectives were selected in this drill-and-practice training game. Therefore, based on previous results from the related works cited in this study, the following research questions were elaborated to guide the present investigation:

1. RQ1: Is there a difference in learning outcomes between male and female students using MatematicaST?
2. RQ2: Are there differences in motivation in terms of interest (RQ2a), competence (RQ2b), and autonomy (RQ2c) between male and female students using MatematicaST?
3. RQ3: Are there differences in perceptions of effort(RQ3a), no pressure (RQ3b), and value (RQ3c) between male and female students using MatematicaST?
4. RQ4: Are there any differences in the effects by gender between the 2 feedback conditions (MTF and STF)?

3.2. MatematicaST Game

"MatematicaST" is a web platform with mathematical games, developed by our research group starting in 2017 as part of an R&D project. It has been piloted in elementary school courses ranging from 3rd to 6th grade across various schools to practice mathematical topics [47]. For the current intervention, three math mini-games were offered: number identification, number ordering, and money counting (see Figure 1), aligned with the national curriculum for 3rd and 4th grade in Chile.The numbers identification game exercises the place value on the symbolic representation of numbers with units, tens, and hundreds. The numbers-ordering game focuses on arranging four numbers from smallest to largest up to 3 digits. Players can drag the numbers to rearrange them and confirm the obtained sequence. The money-counting game focuses on counting coins of different amounts considering coins with values ranging from 1 to 500 (in Chilean peso – CLP).

On each game, the question appears at the top and can be listened to by pressing the play button. Players have three lives (represented by red hearts) and earn ten points for each correct answer. Total points are displayed at the upper right. The platform also features a list of high scores on the right side.

The game considers 2 versions which vary the type of instructional feedback provided to students. In the STF version of the games, students had one opportunity to answer each exercise. After their response, they received feedback indicating if the answer was correct or incorrect, along with the correct solution in case of a wrong answer. In the MTF version, students have three attempts per exercise without losing a life. After each attempt immediate feedback on their answers' correctness (KR) is given. The correct answer is not revealed even after the last try.

For more details on MatematicaST games and versions please refer to [74].

 doi:10.20944/preprints202404.1186.v1

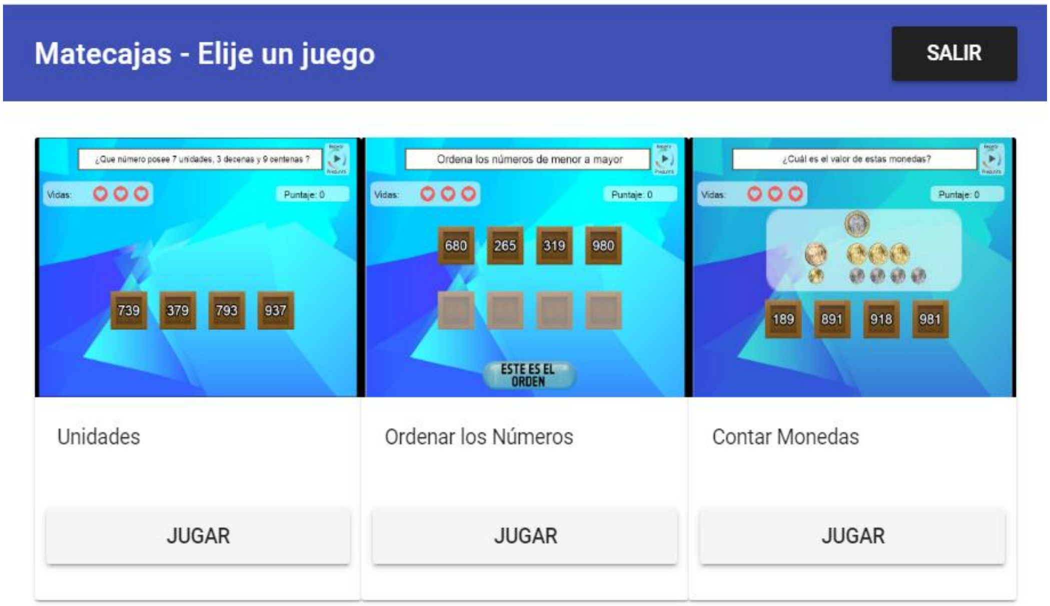

Figure 1. MatematicaST game main menu screen (in Spanish).

3.3. Participants and Design

In this study, students from third and fourth grade from three schools located in the Valparaíso region of Chile were engaged in the intervention. Initially, the study considered the participation of 95 students. However, after controlling for complete participation (student absences on activity days, missing parental consent, or incomplete responses in post-test or post-IMI), the final number of subjects was 81.

The research followed a randomized pre-test-intervention-post-test design, with the gender (male/female) being the first factor of analysis and the feedback type + attempts a second factor (STF/MTF). The STF condition considers a single try with Knowledge of the Correct Response after answering. MTF condition involves a 3-attempts multiple-try with Knowledge of the Result among trials. Students were randomly assigned to these conditions, with 41 participants (23 male, 18 female) in the Multiple-try group and 40 participants (19 male, 21 female) in the STF condition.

3.4. Instruments

3.4.1. Math Tests

Pre and post-tests in this study were meticulously designed based on three specific learning objectives outlined in the Chilean mathematics program and its corresponding guide textbook. Collaborative efforts were made with teachers from participating schools, who actively contributed to the construction of these tests. Their valuable inputs were utilized to validate the instructions, exercises, and testing conditions. Each pre-test and post-test encompassed 18 items, and six exercises related to each learning objective, which included tasks focused on: a) number identification, b) number ordering, and c) money counting. The tests were structured to have a maximum of 9.0 points, ensuring a comprehensive assessment aligned with the targeted learning outcomes.

3.4.2. Motivational Questionnaire

The measurement of motivation following the Self-Determination Theory and its Cognitive Evaluation Theory was operationalized using the Intrinsic Motivation Inventory (IMI) [75]. IMI assesses various constructs related to intrinsic motivation during specific activities and has been employed in numerous studies focusing on intrinsic motivation [13,70,72,74]. This research incorporated six key motivation constructs related to students' self-perceptions of interest,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

competence, pressure, effort, choice (autonomy), and value. The perception tests consisted of 16 sentences rated on a 5-point Likert scale, ranging from 1 ("strongly disagree") to 5 ("strongly agree").

3.5. Procedure

The study was conducted during two school days on math lessons. On day 1, students were initially given a motivational talk outlining the experiment's objectives. Clear instructions about the procedures and the tests to be administered were provided (10 minutes). Subsequently, students took the pre-IMI test and the learning pre-test using a paper-and-pencil format for 20 minutes. Following this, students accessed the web platform and engaged in the MatematicaST game activity on school computers for 1 hour. On day 2, students responded to the post-IMI test and the learning post-test (20 min).

4. Results

Data was processed with SPSS software version 29 for the statistical analyses and graphs generation. Partial eta squared (η^2) was used as a measure of effect size, with values of .01, .06, and .14 for small, medium, and large effect sizes correspondingly, provided by Cohen [76].

4.1. Learnings by Gender

First, Table 1 shows the descriptive statistics for pre- and post-test scores per gender and condition. Then, to assess the overall learnings and the possible effects of gender, a 2-way repeated measures ANOVA was performed considering gender as a between-subjects factor and (pre/post) test type as a within-subjects factor. Results show a difference in pre and post-test scores with F(1, 79)=19.380, p<.001, η^2 =.197, meaning significant learning gains with the activity. Then, the main effects of gender over learnings resulted significant with F(1, 79)=12.678, p<.001, η^2 =.138, indicating differences by gender. When evaluating the effects of gender on test type (pre/post-test) results show that males obtained significantly higher pre-test scores than females (F(1, 79)=11.950, p<.001, η^2 =.131) but not in post-test scores (F(1, 79)=2.926, p=.091, η^2 =.036).

Table 1. Pre-, post-test, and learning gains descriptive statistics per gender and condition.

Gender	Condition	Pre-test		Post-test		Gain		N
		Mean	SD	Mean	SD	Mean	SD	
Female	MTF	6,89	2,16	7,67	1,99	,78	,94	18
	STF	6,99	1,97	7,79	1,55	,80	1,00	21
	Total	6,94	2,03	7,73	1,74	,79	,96	39
Male	MTF	8,54	,80	8,58	,83	,03	,74	23
	STF	7,83	1,53	7,97	1,57	,14	,94	19
	Total	8,22	1,22	8,30	1,25	,08	,82	42
Total	MTF	7,82	1,74	8,18	1,51	,36	,90	41
	STF	7,39	1,80	7,88	1,54	,49	1,01	40
	Total	7,60	1,77	8,03	1,52	,42	,95	81

To consider the pre-test differences, a 2-way ANCOVA was carried out on pre-post-est learning gains, with gender and condition as factors and pre-test as the controlled covariate. It shows a significant single effect of gender on learning gains with F(1, 76)=4.309, p=.041, η^2 =.054, meaning that females learned significantly more than males after adjusting for pre-test scores (See Figure 2). The single effect of the condition was not significant (F(1, 76)=.001, p=.972) meaning that learning gains on both conditions (STF and MTF) were similar.

The 2-way interaction of gender and condition resulted in not being significant (F(1, 76)=.073, p=.788, η^2 =.001). Pairwise comparisons on gender by condition considering Bonferroni adjustment revealed that females learned more than males in both conditions but were not significant with F(1, 76)=1.694, p=.197, η^2 =.022 for MTF and F(1, 76)=2.995, p=.088, η^2 =.038 for STF. Also when analyzing the effect of condition on gender, non-significant differences were obtained for females (F(1, 76)=.027, p=.870) and males (F(1, 76)=.048, p=.828), meaning that learning gains were similar across conditions on males and females.

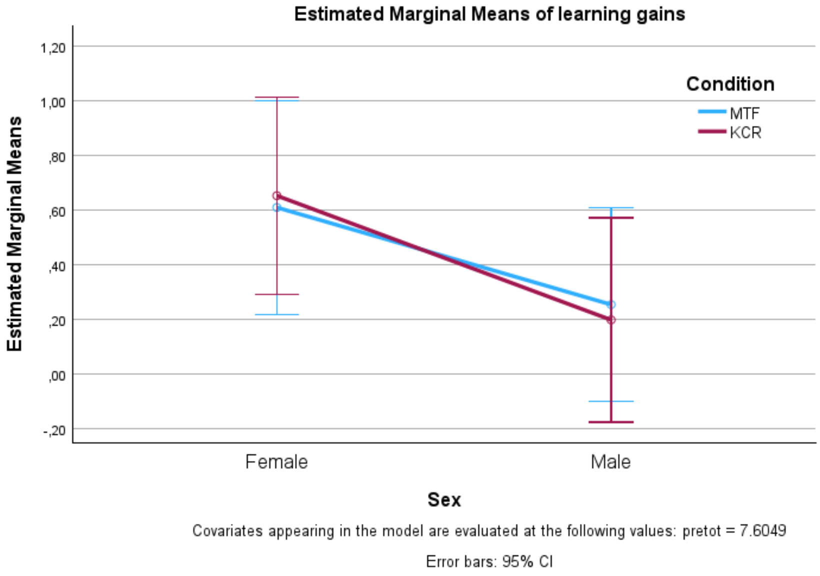

Figure 2. Pre-post learning differences per gender and feedback condition after controlling for pre-test.

4.2. Gender & Motivation

Table 2 shows the descriptive statistics of IMI dimensions per gender and condition. Pre and post-IMI-test values considered measures for the dimensions of interest, perceived competence, perceived choice, effort, no pressure (reverse of pressure), and value. To assess the internal consistency of the questionnaires a Cronbach alpha of .72 for pre-IMI, .69 for post-IMI, and .81 on both perception tests were obtained.

Table 2. IMI-perception constructs' descriptive statistics per condition.

IMI Construct	Gender	Condition	Pre-test		Post-test		Delta		N
			Mean	SD	Mean	SD	Mean	SD	
Interest	Female	MTF	4,64	,48	4,69	,73	,05	,75	18
		STF	4,70	,56	4,85	,38	,15	,69	21
		Total	4,67	,52	4,77	,56	,10	,71	39
	Male	MTF	4,56	,60	4,64	,67	,08	,63	23
		STF	4,58	,59	4,77	,44	,18	,59	19
		Total	4,57	,58	4,70	,57	,13	,60	42
	Total	MTF	4,59	,54	4,66	,69	,07	,68	41
		STF	4,65	,57	4,81	,40	,17	,63	40
		Total	4,62	,55	4,74	,57	,12	,65	81
Competence	Female	MTF	3,96	,83	4,38	,74	,41	,82	18
		STF	4,29	,75	4,29	,92	-,02	,72	21
		Total	4,14	,80	4,33	,83	,18	,79	39
	Male	MTF	4,13	,78	4,56	,69	,42	,82	23
		STF	3,83	,86	4,75	,46	,94	,90	19
		Total	4,00	,82	4,65	,60	,65	,88	42
	Total	MTF	4,06	,80	4,48	,71	,41	,81	41
		STF	4,07	,83	4,51	,77	,44	,93	40
		Total	4,06	,81	4,49	,73	,42	,87	81
Effort	Female	MTF	4,56	,75	4,19	,89	-,36	,74	18
		STF	4,52	,84	4,02	,99	-,50	,87	21
		Total	4,54	,79	4,10	,94	-,44	,80	39

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

	Male	MTF	4,50	,78	3,85	,98	-,65	,99	23
		STF	4,55	,71	4,29	,79	-,26	1,07	19
		Total	4,52	,74	4,05	,92	-,48	1,04	42
	Total	MTF	4,52	,76	4,00	,95	-,52	,89	41
		STF	4,54	,77	4,15	,90	-,39	,96	40
		Total	4,53	,76	4,07	,92	-,46	,93	81
NoPressure	Female	MTF	3,92	,79	3,97	,90	,06	1,26	18
		STF	4,05	1,00	4,40	,90	,36	,91	21
		Total	3,99	,90	4,21	,92	,22	1,08	39
	Male	MTF	3,67	,91	3,83	1,24	,15	1,21	23
		STF	3,71	,73	4,13	1,10	,42	,98	19
		Total	3,69	,83	3,96	1,18	,27	1,11	42
	Total	MTF	3,78	,86	3,89	1,09	,11	1,22	41
		STF	3,89	,89	4,27	1,00	,39	,93	40
		Total	3,83	,87	4,08	1,06	,25	1,09	81
Choice	Female	MTF	3,96	,90	3,99	,66	,02	,92	18
		STF	3,70	1,10	4,10	,75	,39	,97	21
		Total	3,82	1,01	4,05	,70	,22	,95	39
	Male	MTF	3,34	1,16	4,12	,75	,77	,95	23
		STF	3,44	,89	3,99	,95	,55	1,03	19
		Total	3,38	1,04	4,06	,84	,67	,98	42
	Total	MTF	3,61	1,09	4,07	,70	,44	1,00	41
		STF	3,57	1,00	4,05	,84	,46	,99	40
		Total	3,59	1,04	4,06	,77	,45	,98	81
Value	Female	MTF	4,52	,62	4,42	,88	-,11	,69	18
		STF	4,63	,60	4,79	,37	,16	,65	21
		Total	4,58	,60	4,62	,67	,04	,67	39
	Male	MTF	4,46	,52	4,74	,47	,28	,53	23
		STF	4,34	,50	4,50	,94	,16	1,04	19
		Total	4,41	,51	4,63	,72	,22	,80	42
	Total	MTF	4,49	,56	4,60	,69	,11	,63	41
		STF	4,49	,57	4,65	,71	,16	,85	40
		Total	4,49	,56	4,62	,70	,13	,74	81

The differences among post-IMI and pre-IMI values were analyzed. A 3-way repeated measures ANOVA was carried out to evaluate the effects of conditions (MTF / STF), gender (male/female), and the six IMI dimensions on the pre-post IMI delta values.

First, results show a significant single main effect of gender on IMI delta values with $F(1, 77) = 4.053$, $p=.048$, $\eta^2 = .050$, with mean = .253 (95% CI [.113, .392]) for males and mean = .050 (95% CI [-.095, .194]) for females, meaning that males presented higher gains in the overall IMI scores when considering all six constructs related to motivation as a whole.

The interaction among IMI dimensions and gender resulted in non-significant ($F(5, 73) = 1.648$, $p=.158$, $\eta^2 = .101$), the same as the interaction among IMI dimensions and condition ($F(5, 73) = .159$, $p=.977$, $\eta^2 = .011$). However, the 3-way interaction among the 6 IMI dimensions, gender, and condition resulted significant with $F(5, 73) = 2.969$, $p=.017$, $\eta^2 = .169$).

Post-hoc tests considered Bonferroni adjustment for multiple comparisons. When analyzing the simple main effect of gender on IMI dimensions we found significant differences in competence ($F(1, 77) = 7.195$, $p=.009$, $\eta^2 = .085$) and choice ($F(1, 77) = 4.407$, $p=.039$, $\eta^2 = .054$ favoring males over females as Figure 3 shows. For the rest of the IMI dimensions, no significant differences were obtained.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

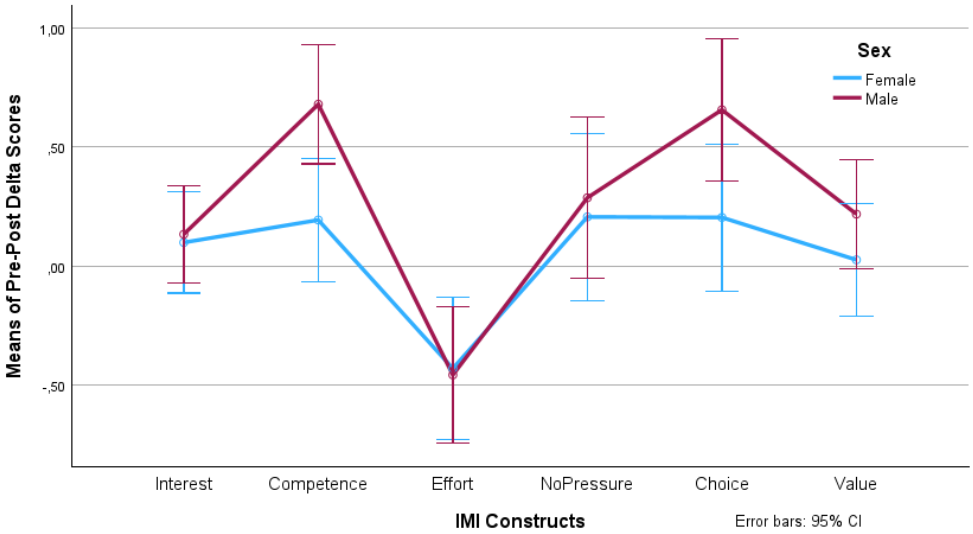

Figure 3. Pre-post delta values of motivational constructs per gender.

Now, when analyzing the 3-way interaction among gender, IMI dimensions, and condition, we obtained that under the MTF condition (see Figure 4) there were no significant differences of gender for the competence IMI-construct (F(1, 77) = .002, p=.963) while for the choice dimension males presented a significantly higher gain than females (F(1, 77) = 5.982, p=.017, η^2 = .072). On the contrary, under the STF condition (see Figure 5) males obtained significantly higher gains for competence (F(1, 77) = 13.956, p<.001, η^2 = .153) while no relevant differences existed among males and females for perceived choice (F(1, 77) = .280, p=.598).

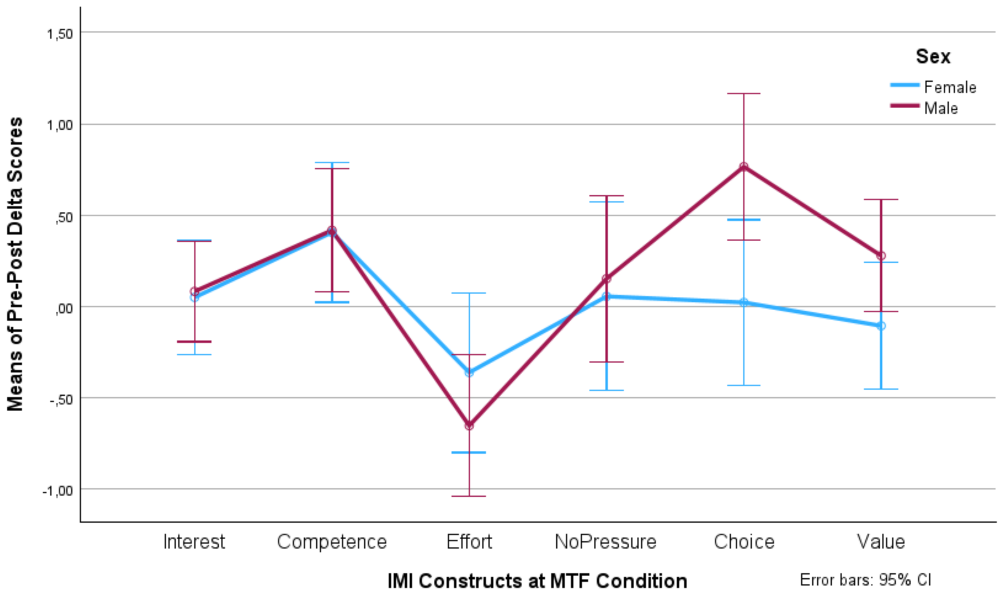

Figure 4. Pre-post delta values of motivational constructs per gender for the MTF condition.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

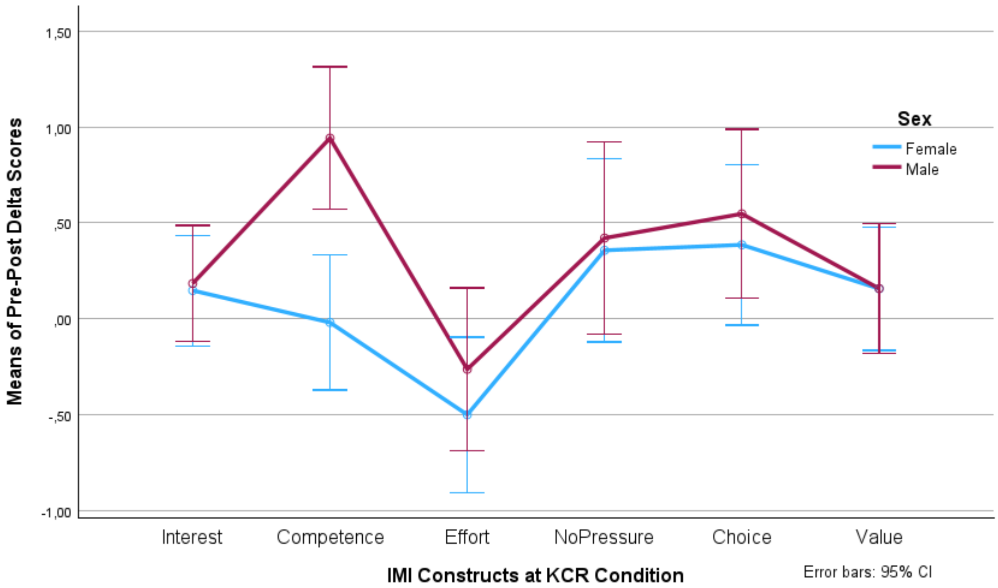

Figure 5. Pre-post delta values of motivational constructs per gender for the STF condition.

5. Discussion

5.1. RQ1: Is There a Difference in Learning Outcomes between Male and Female Students Using MatematicaST?

When analyzing RQ1, results by gender are not similar in terms of learning gains. Females presented higher learning gains than males when adjusting for pre-test scores. At first, our findings support past research suggesting that females present more engagement and learning gains than males under DGBL [7,14,26] and contradict research finding no significant gender differences in learning [16,19].

However, it is relevant to consider that males had higher pre-test scores than females and that post-test scores were similar across genders. Hence, we can think that females outperforming learning gains were at least partially supported by their lower prior knowledge level, as diverse literature states the idea that lower levels of prior knowledge provide more space for learning improvements [31,40,44,77].

Now, from the perspective of the instructional feedback conditions provided in the present experiment, we have that MTF and STF game versions provided similar results in terms of learning gains for each gender. Therefore, the reported gender effects on learning seem not to have been affected by these instructional feedback conditions.

Furthermore, the present results show no learning differences among the two instructional conditions involved in this experiment, namely STF and MTF. Such contradicts past research supporting the idea of MTF outperforming STF in high-order or complex learning outcomes that are not just verbatim tasks [39,50,63,74].

A possible factor that could have affected the results is the presence of an overall high pre-test score (and hence a high prior knowledge level) in general when contrasted to other literature. While past literature usually presents levels of pre-test in the range of 50% to 80% [48,51–53], in our case the average score in the pre-test was 84% and in males under the MTF game version was 95%, severely affecting the possibilities of learning improvements for such a group

5.2. RQ2: Are There Differences in Motivation in Terms Of Interest (RQ2a), Competence (RQ2b), and Autonomy (RQ2c) between Male and Female Students Using MatematicaST?

When looking at gender perception in terms of interest (RQ2a), no differences were assessed among genders. Results support Manero et al. [29], finding no interest differences by gender, and partially contradicting Chung & Chang's study [16], stating that "in a moderate genre digital game,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

female learners' motivation is significantly higher than that of male learners". This is because, from the perspective of learning, our game does appear to have had a moderate gender influence, while in terms of motivation, it does not. This could be due to the basic mechanics of the game not requiring much coordination skill (as in a shooting or jumping platform game). This may have benefited females as evideced by [3,16] they prefer exploration mechanics over competitive ones. Furthermore, the gamification elements used in our MatematicaST game were points, ranking, and levels, which foster progression and competition dynamics. These elements and their associated mechanics could have negatively affected females due to the absence of other game elements such as challenges, storylines, and fun interaction found in Chung & Chang's game [16]. On the other hand, despite the general belief that males find games more relaxing and engaging than females [22], our research revealed that males did not perceive MatematicaST as more interesting. This discrepancy may be attributed to boys' preferences for action games, as highlighted by Khan et al. [14], whereas MatematicaST employs a point-and-click mechanic.

Regarding competence (RQ2b) and autonomy (choice – RQ2c) dimensions, both genders reported positive increments. However, males showed significantly higher increments compared to females on both constructs. Despite females' higher learning gains, males presented higher perceived competence and autonomy, both with medium effect sizes. Such results do not follow the general idea proposed by SDT-CET theory regarding that competence and autonomy are strong predictors of motivation and learning [65–68]. These contradicting results may be attributed to the Dunning–Kruger effect [77], where students tend to significantly overestimate their performance, fostering a belief in their adequate knowledge of a given topic.

Now, from the perspective of females, despite having higher learning gains, they perceived themselves as less competent compared to males. Such results are coherent with Goetz et al. [37] regarding "students' beliefs about their competence in mathematics, with female students reporting lower perceived competence than male despite having the same average grades.”

As competence and autonomy are strong predictors of intrinsic motivation, we can infer that males felt significantly more motivated by the activity. This is corroborated by an overall higher IMI score of males against females when considering all 6 motivational constructs together and with almost a medium effect size. Perhaps such motivation was due to the gaming itself and that males tend to be more competitive than females at that age, especially under a computer gameplay scenario. Also, the point's mechanism and the leaderboard could have positively affected males' motivation.

Our findings support Bonanno & Kommers [22], reporting that boys consider playing games a memorable and unique experience. Also, males' high pre-test and post-test scores could make them feel confident in their knowledge of the math topics involved, increasing their perception of competence and autonomy while reducing their interest in a subject that they feel already mastered.

5.3. RQ3: Are There Differences in Perceptions of Effort(RQ3a), No Pressure (RQ3b), and Value (RQ3c) between Male and Female Students Using MatematicaST?

Concerning the effort construct (RQ3a), no significant differences were observed between males and females. Also, both genders showed a decreasing pre-post delta, which is indicative that students ended up needing less work than they were willing to exert at the beginning of the activity. Although the pre-post test difference is negative, the post-test level remains high (80%), providing indications of the non-existence of factors perceived as controlling by the students[65], which could have negatively affected motivation and learning.

An interesting situation is that females showed a decrease in effort despite perceiving themselves as less competent when compared to males and outperforming them in learning gains.

Regarding the no-pressure construct (RQ3b), no significant differences existed among males and females. By considering pressure and anxiety as related constructs, results support findings from gender math anxiety studies [37], indicating that no gender differences exist for state anxiety, that is, anxiety during specific activities, such in this case, participating in a math training activity with a digital game. Both genders obtained a decrease in their effort and an increase in the no-pressure

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 April 2024 doi:10.20944/preprints202404.1186.v1

dimensions, corroborating the fact that both intervention conditions were suitable to promote learning [65].

On the value dimension (RQ3c), although there were no significant differences across genders, males presented higher levels than females. This is aligned with the idea that for females games are not a unique experience, just another learning alternative, awarding less usefulness than males [22].

5.4. RQ4: Are There Any Differences in the Effects by Gender between the 2 Feedback Conditions (MTF and STF)?

As indicated earlier, there were no learning differences between the two instructional conditions, MTF and STF, both overall and when evaluated for each gender. However, upon examining potential effects on motivational variables resulting from the interaction of gender and instructional condition, the results revealed significant differences. Specifically, it was found that in the MTF condition, there were no significant gender differences regarding the sense of competence, but there were differences in terms of autonomy. Conversely, in the STF condition, gender differences were present for competence but not for the autonomy construct.

These results are relevant as they suggest a possible gender difference between the two instructional implementations (MTF and STF) used in our experiment. They directly indicate that when using multiple attempts, females felt as competent as men in the exercise activity, but at the same time, they felt less free or autonomous than males. One possible explanation could stem from the fact that females perceive themselves as less competent in math despite achieving similar learning outcomes [37]. Then, the mechanic of having multiple attempts (MTF) to answer the same exercise makes the student face an item they know they have answered incorrectly, generating a degree of frustration [55,60,61] and reinforcing the idea of limited effectiveness in an item that needs repetition. If we add to this the fact that the pre-test level (and therefore, prior knowledge) was lower in females, we can assume that they probably used the retry mechanism more often than men when not answering correctly on the first attempt. This aligns with the idea that low achievers tend to feel more frustrated when using MTF [61–63].

As age is an important personal trait in educational contexts [40,44], it could affect multiple attempts' effectiveness and the gender differences found in this work. Therefore, future studies should also include preschoolers [52] and the elderly, with games aiding their cognitive and emotional needs [79].

Likewise, these results open up the possibility of questioning whether there might be other instructional elements and combinations thereof that also generate gender differences, either in terms of learning or motivation and emotions. From this, it becomes necessary to move forward with further studies that first allow for replicating these results and then proceed to evaluate under which other factors and interactions different effects are generated in both genders when varying the number of attempts.

6. Conclusions

The present study provided empirical results on the different effects obtained by gender when using MatematicaST, a digital math training game for primary students. Results indicate that females outperformed males in learning gains, despite presenting lower motivation levels, especially regarding self-perceptions of competence and autonomy. In this way, the present research contributes to the literature by not only providing a successful case of game-based learning favoring females in terms of learning but also by giving insights into the internal emotional processes that could have affected such differences in learning.

This study is perhaps the first research providing empirical evidence on the effects of multiple attempts by gender. An interaction effect between gender and feedback condition was found. Females under multiple-try feedback presented lower autonomy levels than males while in STF there were no differences. On the contrary, females under MTF presented similar competence levels as males but under STF females presented lower levels. Such findings suggest that not all feedback types are gender-neutral, especially the ones involving multiple attempts. It will be relevant to check if those findings extend to other learning contexts and to delve deeper into the factors that enable the generation of such gender differences.

Practical implications regard the use of multiple-try not only in automated assessments or testing but also in game-based formative activities. The use of multiple attempts in existing LMS (e.g., Moodle) activities will generate learning gains in students, especially in females. Also, the design of future educational computer systems should include multiple attempts but take care of the possible gender differences leveraged in this study. Work limitations arise from the nature of the learning objectives used. Although they were of high complexity, future research should include LOs from other mathematical domains. Differences in students' socio-economic status (SES) can be considered other sources of possible limitation. The experiment involved private (higher SES) and subsidized schools (medium SES). However, students from public schools (low SES) were missing so should be considered in the next studies.

Future research directions should also include multiple-try implementations with various attempt numbers, instructional feedback types, and game mechanics. Also, possible gender differences in multiple-try use and its outcomes under all these variants need further study.

Author Contributions: Conceptualization, C.C. and S.R.; methodology, C.C., and S.R.; software, C.C.; validation, C.C., S.R., D.C.-P. and R.M.V.; formal analysis, C.C.; investigation, C.C.; resources, C.C.; data curation, C.C.; writing—original draft preparation, C.C.; writing—review and editing, C.C., D.C.-P. and R.M.V.; visualization, S.R.; supervision, S.R.; project administration, C.C.; funding acquisition, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Pontifical Catholic University of Valparaíso grant number DI 039.360/2023.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Bioethics and Biosecurity Committee of Pontificia Universidad Católica de Valparaíso (protocol code BIOEPUCV-H 659-2023 on 29 June 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available as it may present identification data of the study subjects.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Higgins, K., Huscroft-Dangelo, J., Crawford, L. (2019). Effects of Technology in Mathematics on Achievement, Motivation, and Attitude: A Meta-Analysis. Journal of Educational Computing Research 57(2), 283-319. https://doi.org/10.1177/0735633117748416.
2. Ahuja, M. K., & Thatcher, J. B. (2005). Moving beyond intentions and toward the theory of trying: Effects of work environment and gender on post-adoption information technology use. MIS Quarterly, 29(3), 427–459.
3. Denden, M., Tlili, M., Essalmi, F., Jemni, M., Chen, N.-S., Burgos, D. Effects of gender and personality differences on students’ perception of game design elements in educational gamification, International Journal of Human-Computer Studies, 2021, doi: https://doi.org/10.1016/j.ijhcs.2021.102674
4. Park, C., Kim, D. G., Cho, S., & Han, H. J. (2019). Adoption of multimedia technology for learning and gender difference. Computers in Human Behavior, 92, 288-296.
5. Malik, A., Hiekkanen, K., Hussain, Z., Hamari, J., & Johri, A. (2019). How players across gender and age experience Pokémon Go?. Universal Access in the Information Society, 1-14.
6. Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly, 36(1), 157–178,
7. Lukosch, H. K., Kurapati, S., Groen, D., & Verbraeck, A. (2017). Gender and Cultural Differences in Game-Based Learning Experiences. Electronic Journal of e-Learning 15(4), 310-319.
8. Naglieri, J. A., & Rojahn, J. (2001). Gender differences in planning, attention, simultaneous, and successive (PASS) cognitive processes and achievement. Journal of Educational Psychology, 93(2), p. 430-437.

9. Boghi, A., Rasetti, R., Avidano, F., Manzone, C., Orsi, L., D'agata, F., Caroppo, P., Bergui, M., Rocca, P., Pulvirenti, L., Bradac, G.B., Bogetto, F. Mutani, R. & Mortara, P. (2006). The effect of gender on planning: An fMRI study using the Tower of London task. Neuroimage, 33(3), pp. 999-1010.
10. Bergdahl, N., & Nouri, J. (2020). Student engagement and disengagement in TEL – The role of gaming, gender and non-native students. Research in Learning Technology, 28. https://doi.org/10.25304/rlt.v28.2293.
11. Bergdahl, N., Nouri, J., Fors, U., Knutsson, O. (2019) Engagement, disengagement and performance when learning with technologies in upper secondary school. Computers & Education, 149, Article 103783. https://doi.org/10.1016/j.compedu.2019.103783
12. Jabbar, A. I. A., & Felicia, P. (2015). Gameplay Engagement and Learning in Game-Based Learning: A Systematic Review. Review of Educational Research, 85(4), 740–779. https://doi.org/10.3102/0034654315577210.
13. Chen, C., Law, V., & Huang, K. (2019). The roles of engagement and competition on learner's performance and motivation in game-based science learning. Educational Technology Research and Development, 67(4), 1003-1024. https://doi.org/10.1007/s11423-019-09670-7.
14. Khan, A., Ahmad, F. & 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. Educ Inf Technol, 22, 2767–2804. https://doi.org/10.1007/s10639-017-9622-1.
15. Bahar, M. & Asil, M. (2018). Attitude towards e-assessment: influence of gender, computer usage and level of education, Open Learning. The Journal of Open, Distance and e-Learning, 33(3), 221-237. https://doi.org/10.1080/02680513.2018.1503529.
16. 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. https://doi.org/10.12973/eurasia.2017.01227a.
17. Erturan, S., & Jansen, B. (2015). An investigation of boys' and girls' emotional experience of math, their math performance, and the relation between these variables. Eur. J. Psychol. Educ., 30, 421–435. https://doi.org/10.1007/s10212-015-0248-7.
18. Wang, L. (2020). Mediation Relationships Among Gender, Spatial Ability, Math Anxiety, and Math Achievement. Educ Psychol Rev, 32, 1–15. https://doi.org/10.1007/s10648-019-09487-z.
19. Dorji, U., Panjaburee, P. & Srisawasdi, N. (2015). Gender differences in students' learning achievements and awareness through residence energy saving game-based inquiry playing. Journal of Computers in Education, 2, 227–243. https://doi.org/10.1007/s40692-015-0033-2.
20. Twenge, J. M., & Martin, G. N. (2020). Gender differences in associations between digital media use and psychological well-being: Evidence from three large datasets. Journal of Adolescence, 79, 91–102. https://doi.org/10.1016/j.adolescence.2019.12.018.
21. Wang, Y. S., Wu, M. C., & Wang, H. Y. (2009). Investigating the determinants and age and gender differences in the acceptance of mobile learning. British Journal of Educational Technology, 40(1), 92–118.
22. Bonanno, P., & Kommers, P. A. M. (2008). Exploring the influence of gender and gaming competence on attitudes towards using instructional games. British Journal of Educational Technology, 39(1), 97-109. https://doi.org/10.1111/j.1467-8535.2007.00732.x.
23. Carvalho, R. G. G. (2016). Gender differences in academic achievement: The mediating role of personality. Personality and Individual Differences, 94, 54–58.
24. Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. Journal of Educational Psychology, 101(3), 689.
25. Cipriani, G. P. (2018). Gender difference in willingness to guess after a failure. The Journal of Economic Education, 49(4), 299-306. https://doi.org/10.1080/00220485.2018.1500958.
26. Chang, M., Evans, M., Kim, S., Deater-Deckard K., & Norton, A. (2014). Educational Video Games and Students' Game Engagement. 2014 International Conference on Information Science & Applications (ICISA), Seoul, 2014, pp. 1-3, https://doi.org/10.1109/ICISA.2014.6847390.
27. 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 and Educational Technology, 21(2), 295–303. https://doi.org/10.1007/s10956-011-9319-y.

28. Lester, J. C., Spires, H. A., Nietfeld, J. L., Minogue, J., Mott, B.W., & Lobene, E. V. (2014). Designing gamebased learning environments for elementary science education: A narrative-centered learning perspective. Information Sciences, 264, 4–18. https://doi.org/10.1016/j.ins.2013.09.005.
29. Manero, B., Torrente, J., Fernández-Vara, C. & Fernández-Manjón, B. (2017). Investigating the Impact of Gaming Habits, Gender, and Age on the Effectiveness of an Educational Video Game: An Exploratory Study. IEEE Transactions on Learning Technologies, 10(2), 236-246. https://doi.org/10.1109/TLT.2016.2572702.
30. Jones, R. (2009). Student Engagement Teacher Handbook. Rexford: International Center for Leadership in Education.
31. Yeo, JH., Cho, IH., Hwang, GH. et al. Impact of gender and prior knowledge on learning performance and motivation in a digital game-based learning biology course. Education Tech Research Dev 70, 989–1008 (2022). https://doi.org/10.1007/s11423-022-10099-8
32. Keller, J. M. (1999). Using the ARCS motivational process in computer-based instruction and distance education. New Directions for Teaching and Learning, 78, 39–47. https:// doi. org/ 10. 1002/ tl. 7804
33. Pérez-Fuentes, M. D. C., Núñez, A. , Molero, M. D. M., Gázquez, J. J., Rosário, P., and Núñez, J. C. (2020). The Role of Anxiety in the Relationship between Self-efficacy and Math Achievement. Psicología Educativa, 26(2), 137 - 143. https://doi.org/10.5093/psed2020a7.
34. Namkung, J. M., Peng, P., and Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: a meta-analysis. Rev. Educ. Res. 89, 459–496. https://doi.org/10.3102/0034654319843494.
35. Vanbecelaere, S., Van den Berghe, K., Cornillie, F., Sasanguie, D., Reynvoet, B., & Depaepe, F. (2020). The effects of two digital educational games on cognitive and non-cognitive math and reading outcomes. Computers & Education, 143, Article 103680. https://doi.org/10.1016/j.compedu.2019.103680.
36. Abín, A., Núñez J. C., Rodríguez C., Cueli M., García, T., & Rosário, P. (2020). Predicting Mathematics Achievement in Secondary Education: The Role of Cognitive, Motivational, and Emotional Variables. Front. Psychol., 11, 876. https://doi.org/10.3389/fpsyg.2020.00876.
37. Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do Girls Really Experience More Anxiety in Mathematics?. Psychological Science, 24(10), 2079–2087. https://doi.org/10.1177/0956797613486989.
38. Barana, A.; Marchisio, M.; Sacchet, M. Interactive Feedback for Learning Mathematics in a Digital Learning Environment. *Educ. Sci.* 2021, 11, 279. https://doi.org/10.3390/educsci11060279
39. Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes: A Meta-Analysis. Review of Educational Research, 85(4), 475–511. https://doi.org/10.3102/0034654314564881.
40. Shute, V. J. (2008). Focus on Formative Feedback. Review of Educational Research, 78(1), 153–189. https://doi.org/10.3102/0034654307313795.
41. Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77, 81–112. https://doi.org/10.3102/003465430298487.
42. Candel, C, Vidal-Abarca, E, Cerdán, R, Lippmann, M, Narciss, S. (2020). Effects of timing of formative feedback in computer-assisted learning environments. J Comput Assist Learn, 36, 718– 728. https://doi.org/10.1111/jcal.12439.
43. Andre, T., & Thieman, A. (1988). Level of adjunct question, type of feedback, and learning concepts by reading. Contemporary Educational Psychology, 13(3), 296–307. https://doi.org/10.1016/0361-476X(88)90028-8.
44. Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merril, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), Handbook of research on educational communications and technology (3rd ed., pp. 125–144). Mahwah, NJ: Lawrence Erlbaum.
45. Wang, Z., Gong, S-Y., Xu, S, & Hu, X-E. (2019). Elaborated feedback and learning: Examining cognitive and motivational influences. Computers & Education, 136, 130-140. https://doi.org/10.1016/j.compedu.2019.04.003.
46. Tsai, F.-H., Tsai, C.-C., & Lin, K.-Y. (2015). The evaluation of different gaming modes and feedback types on game-based formative assessment in an online learning environment. Computers & Education, 81, 259–269. https://doi.org/10.1016/J.COMPEDU.2014.10.013.
47. Montazeri, M. & Salimi, E.-A. (2019). Assessing motivation to speak (MTS) and willingness to communicate through metalinguistic corrective feedback. Learning and Motivation 68, 101594.

48. Attali, Y. (2015). Effects of multiple-try feedback and question type during mathematics problem solving on performance in similar problems. Computers & Education, 86, 260-267. https://doi.org/10.1016/j.compedu.2015.08.011.
49. Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. Review of Educational Research, 61(2), 213–238. https://doi.org/10.3102/00346543061002213
50. Clariana, R.B., & Koul, R. (2005). Multiple-Try Feedback and Higher-Order Learning Outcomes. International journal of instructional media, 32(3), 239. https://www.learntechlib.org/p/63335/.
51. Stevenson, C. E. & Hickendorff, M. (2018). Learning to solve figural matrix analogies: The paths children take. Learning and Individual Differences, 66, 16-28. https://doi.org/10.1016/j.lindif.2018.04.010.
52. Agina, A., Kommers, P., & Heylen, D. (2016). Towards Understanding Human-Media Interaction: The Effect of the Computer's Answer-Until- Correct (AUC) vs. Knowledge-of-Result (KR) Task Feedback on Young Users' Behavioral Interaction Development through a Digital Playground. Journal of Childhood & Developmental Disorders, 2(1). https://doi.org/10.4172/2472-1786.100011.
53. Fyfe, E. R. (2016). Providing feedback on computer-based algebra homework in middle-school classrooms. Computers in Human Behavior, 63, 568–574. https://doi.org/10.1016/J.CHB.2016.05.082
54. Clariana, R.B.;Wagner, D.; Murphy, L.C.R. Applying a Connectionist Description of Feedback Timing. *Educ. Technol. Res. Dev.* **2000**, 48, 5–21. http://www.jstor.org/stable/30220265
55. Clariana, R.B., Ross, S.M. & Morrison, G.R. The effects of different feedback strategies using computer-administered multiple-choice questions as instruction. ETR&D 39, 5–17 (1991). https://doi.org/10.1007/BF02298149
56. Guthrie, E. R. (1935). The psychology of learning. New York: Harper & Brothers.
57. Dempsey, J. V. (1993). Interactive Instruction and Feedback. New Jersey: Educational Technology.
58. Noonan, J. V. (1984). Feedback procedures in computer-assisted instruction: Knowledge-of- results, knowledge-of-correct-response, process explanations, and second attempts after ex- planations. Dissertation Abstracts International, 45(1), 131.
59. Schimmel, B. J. (1986). Feedback use by low-ability students in computer-based education. Dissertation Abstracts International, 47(11), 4068
60. Dick, W., Latta, R. Comparative effects of ability and presentation mode in computer-assisted instruction and programed instruction. ECTJ 18, 33–45 (1970). https://doi.org/10.1007/BF02769441
61. Clariana, R. B., & Koul, R. (2006). The effects of different forms of feedback on fuzzy and verbatim memory of science principles. British Journal of Educational Psychology, 76, 259-270. doi:10.1348/000709905X39134
62. Morrison, G. M., Ross, S. M., Gopalakrishnan, M., & Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. Contemporary Educational Psychology, 20, 32–50.
63. Clariana, R. B., & Lee, D. (2001). The effects of recognition and recall study tasks with feedback in a computer-based vocabulary lesson. Educational Technology Research & Development, 49, 23–36. https://doi.org/10.1007/BF02504913.
64. Salomon, G., & Globerson, T. (1987). Skill may not be enough: the role of mindfulness in learning and transfer. International Journal of Educational Research, 11, 623-637.
65. Ryan, R. M. & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. Contemporary Educational Psychology, 61, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860.
66. Ryan, R. & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55, 68-78. https://doi.org/10.1037/0003-066X.55.1.68.
67. Deci, E. L. & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. Plenum Press: New York.
68. Deci, E. & Ryan, R. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. Psychological Inquiry, 11(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01.
69. Liao, C.W., Chen, C.H., & Shih, S.J. (2019). The interactivity of video and collaboration for learning achievement, intrinsic motivation, cognitive load, and behavior patterns in a digital game-based learning environment. Computers & Education, 133, 43-55. https://doi.org/10.1016/j.compedu.2019.01.013.

70. Chen, C.-H., Law, V. (2016). Scaffolding individual and collaborative game-based learning in learning performance and intrinsic motivation. Computers in Human Behavior, 55B, 1201-1212. https://doi.org/10.1016/j.chb.2015.03.010.
71. Huang, Y., Backman, S.J., Backman, K.F., McGuire, F.A., & Moore, D. (2019). An investigation of motivation and experience in virtual learning environments: a self-determination theory. Educ Inf Technol 24, 591–611. https://doi.org/10.1007/s10639-018-9784-5.
72. Proulx, J.-N., Romero, M., & Arnab, S. (2017). Learning Mechanics and Game Mechanics Under the Perspective of Self-Determination Theory to Foster Motivation in Digital Game Based Learning. Simulation & Gaming, 48(1), 81–97. https://doi.org/10.1177/1046878116674399.
73. Liu, Y. C., Wang, W-T, & Lee, T-L. (2020) An Integrated View of Information Feedback, Game Quality, and Autonomous Motivation for Evaluating Game-Based Learning Effectiveness. Journal of Educational Computing Research, https://doi.org/10.1177/0735633120952044
74. Cubillos, C.; Roncagliolo, S.; Cabrera-Paniagua, D. Learning and Motivation When Using Multiple-Try in a Digital Game for Primary Students in Chile. *Educ. Sci.* **2023**, *13*, 1119. https://doi.org/10.3390/educsci13111119
75. Monteiro, V., Mata, L., & Peixoto, F. (2015). Intrinsic Motivation Inventory: Psychometric Properties in the Context of First Language and Mathematics Learning. Psicol. Reflex. Crit. 28 (3). https://doi.org/10.1590/1678-7153.201528302.
76. Cohen, J (1988) Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.
77. Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. Journal of Educational Psychology, 104(4), 1094–1108. https://doi.org/10.1037/a0028389.
78. Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. Journal of Personality and Social Psychology, 77(6), 1121–1134. https:// doi. org/ 10. 1037/ 0022- 3514. 77.6. 1121
79. Rienzo, A., Cubillos, C. (2020). Playability and Player Experience in Digital Games for Elderly: A Systematic Literature Review. Sensors, 20(14), 3958. https://doi.org/10.3390/s20143958.