Playing an Educational Game Featuring Procedural Content Generation: Which Attributes Impact Players' Curiosity?

Luiz Rodrigues - Londrina State University - luiz_rodrigues17@hotmail.com
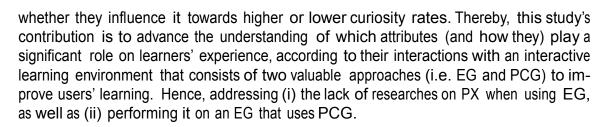Jacques Brancher - Londrina State University - jacques@uel.br

Abstract. *Understanding which and how attributes impact player experience can contribute to designing more tailored tools, providing concerns on how to improve these. However, there is a gap in the understanding of what impacts learners' experience when interacting with Educational Games (EG) featuring Procedural Content generation (PCG) as these have been scantly used together. This article presents an empirical study on which attributes impact both math's and game's curiosity of players when interacting with an EG that uses PCG. The results show the attributes that led to higher or lower curiosity, as well as which of them are associated with it. Hence, advancing the understanding of what drives players' curiosity, contributing to the design of EG that feature PCG.*
Keywords: procedural content generation, educational game, player modeling, curiosity, digital math game.

1. Introduction

Educational Games (EG) can provide players with multiple learning benefits, however, negative outcomes might also emerge according to different situations [Boyle et al. 2016]. A possible reason for the negative outcomes is the lack of focus on Player Experience (PX) since the focus is often on the learning objectives being sought when designing these tools [Eckardt and Robra-Bissantz 2018]. Hence, demonstrating the need for further studies on PX from playing EG. These can reveal different players' groups and how their characteristics affect their experiences and performances [Rodrigues and Brancher 2018, Horn et al. 2016]. Thereby, modeling EG' players can provide valuable insights into how to improve their experience and, consequently, their learning.

Furthermore, developing these games might be a complex, costly and challenging task [Boyle et al. 2016]. Even general purpose games share these problems, often requiring manpower, significant time and budget, and a multidisciplinary team [Hendrikx et al. 2013]. Procedural Content Generation (PCG) algorithms can remedy these problems, automatically creating specific games' parts (e.g. levels or puzzles) and, therefore, aiding and improving the EG' development process [Hendrikx et al. 2013]. However, PCG's use in EG is yet an emerging field of research [Dong and Barnes 2017, Rodrigues et al. 2017]. Consequently, there is a gap in the understanding of which attributes impact PX from interactions with procedurally generated contents in this context. Thus, to expand on this gap, this research investigated learners from an EG based on which attributes impact their experiences in terms of an educational-related component: players' game and math curiosity (e.g., The game motivated me to learn more about math).

Our goal was tackled with an empirical study of the relationship between 16 attributes of two types (i.e. personal and behavioral) and self-reports of players' curiosity, provided after playing an EG featuring PCG. Additionally, this study provides concerns with respect to how these attributes impact the target PX component, demonstrating

whether they influence it towards higher or lower curiosity rates. Thereby, this study's contribution is to advance the understanding of which attributes (and how they) play a significant role on learners' experience, according to their interactions with an interactive learning environment that consists of two valuable approaches (i.e. EG and PCG) to improve users' learning. Hence, addressing (i) the lack of researches on PX when using EG, as well as (ii) performing it on an EG that uses PCG.

2. Literature Review

PCG have has applied in EG approaching a varied set of subjects, such as math [Rodrigues et al. 2017, Smith et al. 2012, Butler et al. 2015, Lara et al. 2018], computing [Horn et al. 2016, Dong and Barnes 2017, Valls-Vargas et al. 2017, Dezani et al. 2017], resource management [Aslam et al. 2017, Luo et al. 2017, Grappiolo et al. 2011], evolution [Soule et al. 2017], therapy [Duval et al. 2017], and reading [Hooshyar et al. 2018]. In those studies, PCG was mostly used to generate levels or its components (e.g. puzzles). Furthermore, most of those research conducted user studies with respect to their interaction with the game. The exceptions are the cases where the algorithms were analyzed [Soule et al. 2017, Smith et al. 2012, Valls-Vargas et al. 2017, Dong and Barnes 2017] or no user evaluation was performed [Dezani et al. 2017]. In contrast, works reporting user-based studies mainly evaluated users' behavior [Butler et al. 2015], affective states' changes [Lara et al. 2018], learning gains [Hooshyar et al. 2018, Rodrigues et al. 2017, Horn et al. 2016], in-game performance [Aslam et al. 2017], opinions [Duval et al. 2017], and the proposed approach's performance [Luo et al. 2017, Grappiolo et al. 2011].

Thereby, showcasing (i) math and computer science, as well as levels, are the topics and content type, respectively, that received more attention from the academic community; and (ii) despite that some of those evaluated PX components, no research evaluated which players' attributes impact on it. Therefore, the understanding of what attributes, and how they affect players when interacting with an EG that features an advanced technique of content generation is yet a gap to the best of the authors' knowledge. Thus, demonstrating the novelty of this research, which can contribute to the design of more tailored game experiences, as well as the development of player models.
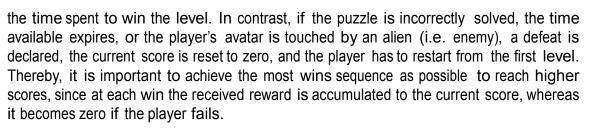
3. Materials and Method

The experiment's goal was to identify which attributes impact on players' short-term curiosity through an empirical study, based on their interaction with an EG featuring PCG.

3.1. Testbed Game

This study's testbed was the EG *SpaceMath*[1], a DMG that fosters the practice of basic math operations. This game was selected by convenience, as we had access to both players' in-game and demographic data (note that all personal information was anonymous) as well as it is free-to-play and might be accessed via web-browsers. In this game, players' goal is to explore as many universes (levels) as possible, wherein solving math puzzles (i.e. basic math operations) is the key to advance. Players advance to the next universe every time they correctly solve the level's puzzle, receiving a score as a reward, based on

---

[1]Available at http://spacemath.rpbtecnologia.com.br

the time spent to win the level. In contrast, if the puzzle is incorrectly solved, the time available expires, or the player's avatar is touched by an alien (i.e. enemy), a defeat is declared, the current score is reset to zero, and the player has to restart from the first level. Thereby, it is important to achieve the most wins sequence as possible to reach higher scores, since at each win the received reward is accumulated to the current score, whereas it becomes zero if the player fails.

To solve levels' math puzzles, players are required to explore the universes. These consist of multiples boxes, spread within them, which hide the pieces (numbers) that solve the puzzles. Consequently, the game makes players explore universes to find puzzles' solutions. We highlight that the game is not intended to teach how to solve the basic arithmetic operations of math. Instead, it aims at leading players to practice solving these tasks as much as possible. Hence, the game's approach is based on repetitive learning [Ke 2008], wherein players are required to constantly solve an educational task. Thereby, it is necessary that the game features large amounts of contents (e.g. levels and math puzzles) to avoid players memorizing their solutions. Thus, the use of PCG to create both its levels and math puzzles are of utmost contribution, allowing the game to provide a pseudo-infinite number of contents with reduced human effort. Therefore, in summary, the game is valuable to education by fostering the practice of a key concept of math, while making players explore game levels to solve math puzzles. Nevertheless, describing the PCG methods is not the scope of this article.

3.2. Players' Attributes

This experiment used two attributes types: personal data and in-game metrics regarding participants' interaction with the game. Table 1 displays all attributes of both types along to a brief description for each one. In summary, participants' attributes that had their impact on curiosity analyzed are (i) personal data that mainly concerns players' demographics, schooling, gaming habits, and affinity to math, and (ii) in-game data that concerns players' performance within the testbed game and the duration of their interaction with it.

3.3. Curiosity Measure

Players' curiosity was measured using an adapted version of the questionnaire introduced by [Wouters et al. 2011]. Aiming to evaluate curiosity in general and specific terms of a serious game, it features seven statements (referred to as C1, C2, and so on) to be completed in the Likert-scale. Although the original scale was seven-point, it was used with a five-point scale in [Moser et al. 2012]. Hence, we also adopted the five-point scale as it is more common for Brazilians. Therefore, participants should choose from "Completely Disagree" to "Completely Agree", encoded from 1 to 5 accordingly. Moreover, two other adaptations had to be made. One emerged from games being of different genres and approaching distinct subjects. The other reason was our participants' native language, which is Portuguese rather than English. Therefore, we translated it to Portuguese and adapted it to state about math and game levels rather than cancer and game world. Despite that, based on this experiment's data, the questionnaire's internal consistency was high (Cronbach's alpha = 0.88); even higher than the original version (0.71). A literal translation (Portuguese to English) of the questions used on this experiment, the Portuguese version, and the original questions as designed by [Wouters et al. 2011] are available at http://bit.ly/2GaSG0s

Table 1. Captured attributes.

Attribute	Description
Age	Participant's current age
Genre	Whether the participant is male or female
Has Net	Whether the participant has a computer with internet access at home
School Type	School type the participant attends (local, public, federal or private)
School Stage	Whether the participant is at elementary, middle or some stage above
Playing Time	Participant's average playing time per week in hours
Gamer	Whether the participant considers itself a gamer
Likes Math	Enjoyment of math in a five-point Likert scale
Knows Math	Self-considered knowledge of math in a five-point Likert scale
Avg. Score	Average score per level
Max. Score	Maximum accumulated score achieved
Avg. Shots	Average of shots fired per level
Avg. Time	Average of time spent to complete each level in seconds
Sum. Time	Total time spent playing the game in seconds
H. Seq	Largest sequence of wins achieved
W. Rate	Ratio of wins per levels played

3.4. Participants

On average, the 265 participants (40% females; 60% males) had 14.5 ($\pm$6.5) years, play 13.3 ($\pm$26.4) hours per week, 47.5% of them considered themselves as gamers, and 91.3% reported to have internet (*HasNet*; see Table 1). Most participants go to private schools (72.5%) and were either finishing elementary school or at the beginning of middle school (68.7%). In five-point Likert-scales, 60% reported the top two scores on how much they like math and 56% reported the middle score (3) on how much they know the subject.

3.5. Procedure

Participants were reached through e-mail lists and via colleagues teachers that agreed to perform the following procedure in some of their classes. Firstly, participants were introduced to the game and to the research. Next, they were asked to register into the game, providing the aforementioned personal data. Then, participants played exactly 20 levels of the testbed to guarantee an equal short-term interaction for everyone. This interaction originated the analyzed in-game data. Lastly, they reported their opinions completing the questionnaire, which the game system automatically administered.

3.6. Data Analysis Process

The data analysis relied on hypothesis tests considering an alpha level of 0.05 for all tests. Three variable types were explored: (i) dichotomous (i.e. binary) are participants' genre and whether they are gamers and have internet; (ii) continuous (i.e. quantitative) are all in-game attributes and participants' age and playing time; and (iii) ordinal (e.g. Likert-scales and variables that can be ordered or ranked) are likes math, knows math, and both school stage and type. To rank the school type variable, we relied on the International Student Assessment Program[2], which evaluates students' knowledge worldwide. Its 2015

[2]Page 173 at https://bit.ly/2oGw7VG

results demonstrate students from Brazilian's federal schools yielded the best proficiency in math, followed by students from private, public and local institutions. Thereby, considering *SpaceMath*'s educational subject, we based the school type's rank on this evidence. Therefore, school type was ranked as follows: federal, 4; private, 3; public, 2; local, 1.

Dichotomous variables' impact was evaluated through Chi-Squared independence tests, with the null hypothesis that variables of this type and curiosity are independent and the alternative hypothesis that those are associated. Ordinal and continuous variables were evaluated through correlation tests, where Kendall's tau-b (τ) test was selected over Pearson's (r) due to the curiosity measures (i.e. ordinal) and over Spearman's based on the recommendation of [Croux and Dehon 2010]. Also, τ coefficients were transformed into r for an easier interpretation, based on [Walker 2003]. Hence, to interpret the results, all correlation analyses considered the following, based on τ coefficients converted to r: strong if greater than 0.5, moderate if between 0.3-0.5, small if 0.1-0.3, and trivial otherwise [Cohen 1977]. Nevertheless, the original τ coefficient and its conversion to r are presented, and correlation significance is based on p-values of Kendall's tests only.

4. Results

This section presents this article's experiment results, showing the correlation tests' results at first. Tables 2, 3, and 4 demonstrate the correlation coefficients from rank, performance, and the remaining non-dichotomous attributes to the seven curiosity statements, respectively. They present both the τ coefficients from the relationship analyses as well as their values converted to r within parentheses.

Table 2. Rank attributes' impact on curiosity based on their correlation. Data represented as τ (r). Significance based on τ test.

Comp	Sc. Stage	Sc. Type	L. Math	K. Math
C1	-0.22 (-0.33)*	-0.06 (-0.09)	0.20 (0.31)*	0.05 (0.08)
C2	-0.12 (-0.19)*	-0.08 (-0.12)	0.22 (0.33)*	0.18 (0.28)*
C3	-0.21 (-0.32)*	-0.03 (-0.04)	0.17 (0.26)*	0.13 (0.2)*
C4	-0.13 (-0.21)*	-0.11 (-0.17)	0.22 (0.35)*	0.14 (0.22)*
C5	-0.05 (-0.08)	-0.05 (-0.08)	0.18 (0.28)*	0.08 (0.12)
C6	-0.17 (-0.26)*	-0.04 (-0.06)	0.16 (0.25)*	0.14 (0.21)*
C7	-0.17 (-0.26)*	-0.04 (-0.06)	0.26 (0.39)*	0.11 (0.18)*
Key: * = $p < 0.05$				

With respect to rank attributes' impact, the findings show significant small to moderate correlations from participants' school stage, liking math, and knowing math to their curiosity. Of these, there are only two exceptions: the relationship from both school stage and knowing math to C5, which are insignificant. Similarly, participants' age has small to moderate significant correlations to all curiosity statements but C5. In contrast, both participants' school type and weekly playing time (WPT) had trivial insignificant correlations to the target variable. With respect to in-game attributes, the findings show that nearly all performance attributes had statistically significant correlations to their curiosity, ranging from small to moderate impacts. The exceptions were the highest sequence of wins and maximum score, which had insignificant relationships to C2, C4, and C5. Differently, time and shot-related attributes were mostly insignificantly correlated to the

Table 3. Performance attributes' impact on curiosity based on their correlation. Data represented as τ (r). Significance based on τ test.

Comp	Avg. Score	H. Seq	W. Rate	Max. Score
C1	-0.20 (-0.30)*	-0.14 (-0.22)*	-0.18 (-0.27)*	-0.15 (-0.23)*
C2	-0.17 (-0.26)*	-0.08 (-0.13)	-0.18 (-0.28)*	-0.09 (-0.14)
C3	-0.22 (-0.34)*	-0.10 (-0.15)*	-0.17 (-0.26)*	-0.13 (-0.20)*
C4	-0.14 (-0.22)*	-0.07 (-0.11)	-0.12 (-0.19)*	-0.09 (-0.14)
C5	-0.15 (-0.23)*	-0.05 (-0.09)	-0.14 (-0.21)*	-0.06 (-0.10)
C6	-0.23 (-0.36)*	-0.15 (-0.24)*	-0.22 (-0.34)*	-0.16 (-0.25)*
C7	-0.17 (-0.27)*	-0.13 (-0.20)*	-0.18 (-0.28)*	-0.14 (-0.21)*
Key: * = $p < 0.05$				

Table 4. Remaining attributes' impact on curiosity based on their correlation. Data represented as τ (r). Significance based on τ test.

Comp	Avg. Time	Avg. Shots	Total Time	Age	WPT
C1	0.05 (0.08)	-0.13 (-0.21)*	0.06 (0.09)	-0.23 (-0.35)*	-0.05 (-0.07)
C2	0.03 (0.04)	-0.04 (-0.07)	0.03 (0.05)	-0.16 (-0.25)*	-0.01 (-0.01)
C3	0.09 (0.15)*	-0.05 (-0.08)	0.10 (0.16)*	-0.23 (-0.36)*	-0.01 (-0.02)
C4	0.07 (0.12)	-0.06 (-0.10)	0.08 (0.12)	-0.15 (-0.23)*	-0.06 (-0.1)
C5	0.04 (0.06)	-0.04 (-0.06)	0.04 (0.07)	-0.07 (-0.12)	-0.02 (-0.04)
C6	0.08 (0.12)	-0.12 (-0.19)*	0.09 (0.13)	-0.18 (-0.27)*	-0.09 (-0.14)
C7	0.05 (0.07)	-0.08 (-0.13)	0.05 (0.08)	-0.19 (-0.29)*	0.03 (0.05)
Key: * = $p < 0.05$; WPT = weekly playing time.					

target variable. The exceptions, in this case, were time-related variables having significant correlations to C3 and average of shots having a significant correlation to both C1 and C6. Nevertheless, all these exceptions presented small correlation coefficients.

Furthermore, Tables 5, 6, and 7 show the association from the following attributes to players' curiosity, respectively: genre, gamer, and has internet. They present the distribution of the questionnaire's answers - 1 to 5 - according to participants' attributes - A - (e.g. male or female) for all curiosity statements - CS. Also, they demonstrate the percentage that the number of answers received represent for each attribute (e.g. 4% of females answered completely disagree - 1). Concerning dichotomous attributes' impact, the results show that, overall, few curiosity statements are associated with them. For participants' genre, no curiosity statement was associated with it. Differently, being a gamer is associated with nearly all curiosity statements, wherein the exceptions are C1 and C3. Lastly, having internet is associated with C2 and C6 only. Thus, while genre is not associated with players' curiosity at all, being a gamer is associated with almost every statement, whereas having internet is associated with only two out of seven statements.

5. Concluding Remarks, Discussion and Future Works

This article presented an empirical study on which attributes impact players' curiosity, considering their self-reported experiences after playing an EG that uses PCG. A series of correlation tests were performed to model these learners, revealing what attributes impacted their curiosity, considering both their personal characteristics and in-game behav-
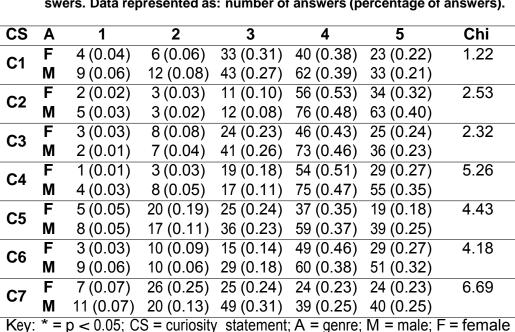
Table 5. Genre impact on curiosity based on its association to participants' answers. Data represented as: number of answers (percentage of answers).

CS	A	1	2	3	4	5	Chi
C1	F	4 (0.04)	6 (0.06)	33 (0.31)	40 (0.38)	23 (0.22)	1.22
	M	9 (0.06)	12 (0.08)	43 (0.27)	62 (0.39)	33 (0.21)	
C2	F	2 (0.02)	3 (0.03)	11 (0.10)	56 (0.53)	34 (0.32)	2.53
	M	5 (0.03)	3 (0.02)	12 (0.08)	76 (0.48)	63 (0.40)	
C3	F	3 (0.03)	8 (0.08)	24 (0.23)	46 (0.43)	25 (0.24)	2.32
	M	2 (0.01)	7 (0.04)	41 (0.26)	73 (0.46)	36 (0.23)	
C4	F	1 (0.01)	3 (0.03)	19 (0.18)	54 (0.51)	29 (0.27)	5.26
	M	4 (0.03)	8 (0.05)	17 (0.11)	75 (0.47)	55 (0.35)	
C5	F	5 (0.05)	20 (0.19)	25 (0.24)	37 (0.35)	19 (0.18)	4.43
	M	8 (0.05)	17 (0.11)	36 (0.23)	59 (0.37)	39 (0.25)	
C6	F	3 (0.03)	10 (0.09)	15 (0.14)	49 (0.46)	29 (0.27)	4.18
	M	9 (0.06)	10 (0.06)	29 (0.18)	60 (0.38)	51 (0.32)	
C7	F	7 (0.07)	26 (0.25)	25 (0.24)	24 (0.23)	24 (0.23)	6.69
	M	11 (0.07)	20 (0.13)	49 (0.31)	39 (0.25)	40 (0.25)	

Key: * = $p < 0.05$; CS = curiosity statement; A = genre; M = male; F = female

Table 6. Gamer impact on curiosity based on its association to participants' answers. Data represented as: number of answers (percentage of answers).

CS	A	1	2	3	4	5	Chi
C1	N	5 (0.04)	12 (0.09)	47 (0.34)	52 (0.37)	23 (0.17)	8.16
	Y	8 (0.06)	6 (0.05)	29 (0.23)	50 (0.40)	33 (0.26)	
C2	N	1 (0.01)	3 (0.02)	14 (0.10)	78 (0.56)	43 (0.31)	9.65*
	Y	6 (0.05)	3 (0.02)	9 (0.07)	54 (0.43)	54 (0.43)	
C3	N	3 (0.02)	7 (0.05)	39 (0.28)	64 (0.46)	26 (0.19)	4.25
	Y	2 (0.02)	8 (0.06)	26 (0.21)	55 (0.44)	35 (0.28)	
C4	N	1 (0.01)	4 (0.03)	25 (0.18)	72 (0.52)	37 (0.27)	10.38*
	Y	4 (0.03)	7 (0.06)	11 (0.09)	57 (0.45)	47 (0.37)	
C5	N	4 (0.03)	22 (0.16)	43 (0.31)	45 (0.32)	25 (0.18)	14.37*
	Y	9 (0.07)	15 (0.12)	18 (0.14)	51 (0.40)	33 (0.26)	
C6	N	3 (0.02)	11 (0.08)	25 (0.18)	68 (0.49)	32 (0.23)	13.30*
	Y	9 (0.07)	9 (0.07)	19 (0.15)	41 (0.33)	48 (0.38)	
C7	N	9 (0.06)	28 (0.20)	47 (0.34)	30 (0.22)	25 (0.18)	10.17*
	Y	9 (0.07)	18 (0.14)	27 (0.21)	33 (0.26)	39 (0.31)	

Key: * = $p < 0.05$; CS = curiosity statement; A = gamer; Y = yes; N = no

ior, and whether those attributes had positive or negative impacts on the target variable, along with the degree (e.g. small or moderate) of these impacts. In addition, independence tests revealed which dichotomous attributes are associated with participants' curiosity.

The findings contribute by showcasing which characteristics impact on PX, in terms of players' math curiosity, when interacting with an EG that uses a scantly explored technique (i.e. PCG) within this context. Thus, advancing the understanding of how interactions with this type of tool are affected, collaborating to different directions. One
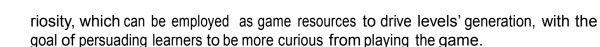
Table 7. *Has Net* **impact on curiosity based on its association to participants' answers. Data represented as: number of answers (percentage of answers).**

CS	A	1	2	3	4	5	Chi
C1	N	1 (0.04)	2 (0.09)	5 (0.22)	8 (0.35)	7 (0.30)	1.69
	Y	12 (0.05)	16 (0.07)	71 (0.29)	94 (0.39)	49 (0.20)	
C2	N	1 (0.04)	4 (0.17)	9 (0.39)	9 (0.39)	1 (0.04)	37.71*
	Y	6 (0.02)	6 (0.02)	19 (0.08)	123 (0.51)	88 (0.36)	
C3	N	2 (0.09)	1 (0.04)	5 (0.22)	10 (0.43)	5 (0.22)	6.37
	Y	3 (0.01)	14 (0.06)	60 (0.25)	109 (0.45)	56 (0.23)	
C4	N	1 (0.04)	1 (0.04)	3 (0.13)	12 (0.52)	6 (0.26)	1.13
	Y	4 (0.02)	10 (0.04)	33 (0.14)	117 (0.48)	78 (0.32)	
C5	N	2 (0.09)	6 (0.26)	4 (0.17)	8 (0.35)	3 (0.13)	4.65
	Y	11 (0.05)	31 (0.13)	57 (0.24)	88 (0.36)	55 (0.23)	
C6	N	3 (0.13)	4 (0.17)	8 (0.35)	8 (0.35)	3 (0.13)	10.67*
	Y	9 (0.04)	20 (0.08)	40 (0.17)	101 (0.42)	72 (0.30)	
C7	N	2 (0.09)	6 (0.26)	5 (0.22)	5 (0.22)	5 (0.22)	1.69
	Y	16 (0.07)	40 (0.17)	69 (0.29)	58 (0.24)	59 (0.24)	

Key: * = $p < 0.05$; CS = curiosity statement; A = *HasNet*; Y = yes; N = no

is that practitioners can exploit these concerns when seeking to tease players' curiosity through these tools. Hence, teachers and other professionals that aim to use EG focused on math can determine what learners will be more curious from playing the game. Consequently, they can select the specific audience that will benefit the most according to their intents. Another direction is that designers/developers can benefit from these insights to design real-time adaptation mechanisms. They can use them to create player models to adjust games towards improving learners' curiosity based on their distinct characteristics, a valuable step towards tackling the challenge of presenting personalized educational games [Santos and Isotani 2018].

Despite that, this study presents some limitations, such as its sample being composed of participants that are unfamiliarized with EG, which might have influenced them toward being more curious than they would be otherwise. Similarly, since participants were aware they were participating in a research, it is possible they wondered that answering the questionnaire positively was expected. Consequently, this might have biased their responses and, therefore, the analyses of what factors impacted their curiosity. Also, there is the fact that most players go to the same school type, wherein a sample with higher variation could lead to different outcomes. Nevertheless, we sought to mitigate these limitations by looking for a representative heterogeneous sample in terms of demographic attributes. However, participants' school did not meet this intent, requiring further research yield more generalizable findings. Thereby, although the limitations that threaten this study, we argue its results are valid based on the representative sample size, the use of a consistent instrument to capture players' curiosity, and data's empirical analyses.

As future research, we suggest investigations on to what extend curiosity can benefit other PX factors (e.g. fun) and players' learning, besides, e.g. leading them to seek for more knowledge with respect to math (i.e. C6). Also, these findings can be used towards the development of computational models to automatically predict players' cu-

riosity, which can be employed as game resources to drive levels' generation, with the goal of persuading learners to be more curious from playing the game.

Acknowledgements

References

Aslam, H., Sidorov, A., Bogomazov, N., Berezyuk, F., and Brown, J. A. (2017). Relief camp manager: A serious game using the world health organization's relief camp guidelines. In Squillero, G. and Sim, K., editors, *Applications of Evolutionary Computation*, pages 407–417, Cham. Springer International Publishing.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., and Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, 94:178 – 192.

Butler, E., Andersen, E., Smith, A. M., Gulwani, S., and Popović, Z. (2015). Automatic game progression design through analysis of solution features. In *Proceedings of the ACM CHI*, pages 2407–2416. ACM.

Cohen, J. (1977). Chapter 3 - the significance of a product moment rs. In Cohen, J., editor, *Statistical Power Analysis for the Behavioral Sciences*, pages 75 – 107. Academic Press.

Croux, C. and Dehon, C. (2010). Influence functions of the spearman and kendall correlation measures. *Statistical Methods & Applications*, 19(4):497–515.

Dezani, H., Nakamura, R., and Marranghello, N. (2017). Automatic generation of interactive narratives for an educational adventure game using petri net. In *Proceedings of SBGames 2017*.

Dong, Y. and Barnes, T. (2017). Evaluation of a template-based puzzle generator for an educational programming game. In *Proceedings of the FDG*, pages 40:1–40:4. ACM.

Duval, J., Rubin, Z., Goldman, E., Antrilli, N., Zhang, Y., Wang, S.-H., and Kurniawan, S. (2017). Designing towards maximum motivation and engagement in an interactive speech therapy game. In *Proceedings of the IDC*, pages 589–594. ACM.

Eckardt, L. and Robra-Bissantz, S. (2018). Playtesting for a better gaming experience: Importance of an iterative design process for educational games research-in-progress. In *Proceedings of the PACIS*.

Grappiolo, C., Cheong, Y., Togelius, J., Khaled, R., and Yannakakis, G. N. (2011). Towards player adaptivity in a serious game for conflict resolution. In *Proceedings of the VS-Games*, pages 192–198.

Hendrikx, M., Meijer, S., Van Der Velden, J., and Iosup, A. (2013). Procedural content generation for games: A survey. *ACM Trans. Multimedia Comput. Commun. Appl.*, 9(1):1:1–1:22.

Hooshyar, D., Yousefi, M., Wang, M., and Lim, H. (2018). A data-driven procedural-content-generation approach for educational games. *Journal of Computer Assisted Learning*, pages 1–9.

Horn, B., Clark, C., Strom, O., Chao, H., Stahl, A. J., Harteveld, C., and Smith, G. (2016). Design insights into the creation and evaluation of a computer science educational game. In *Proceedings of the ACM SIGCSE*, pages 576–581. ACM.

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from game-play? *Computers & Education*, 51(4):1609 – 1620.

Lara, C. A., Mitre-Hernandez, H., Flores, J., and Perez, H. (2018). Induction of emotional states in educational video games through a fuzzy control system. *IEEE Transactions on Affective Computing*, pages 1–12.

Luo, L., Yin, H., Cai, W., Zhong, J., and Lees, M. (2017). Design and evaluation of a data-driven scenario generation framework for game-based training. *IEEE Transactions on Computational Intelligence and AI in Games*, 9(3):213–226.

Moser, C., Fuchsberger, V., and Tscheligi, M. (2012). Rapid assessment of game experiences in public settings. In *Proceedings of the FnG*, pages 73–82. ACM.

Rodrigues, L., Bonidia, R. P., and Brancher, J. D. (2017). A math educational computer game using procedural content generation. In *Proceedings of the SBIE*.

Rodrigues, L. and Brancher, J. D. (2018). Improving players' profiles clustering from game data through feature extraction. In *Proceedings of SBGames 2018*.

Santos, W. O. and Isotani, S. (2018). Desenvolvimento de jogos educativos? desafios, oportunidades e direcionamentos de pesquisa. *RENOTE*, 16(2).

Smith, A. M., Andersen, E., Mateas, M., and Popović, Z. (2012). A case study of expressively constrainable level design automation tools for a puzzle game. In *Proceedings of the FDG*, pages 156–163. ACM.

Soule, T., Heck, S., Haynes, T. E., Wood, N., and Robison, B. D. (2017). Darwin's demons: Does evolution improve the game? In Squillero, G. and Sim, K., editors, *Applications of Evolutionary Computation*, pages 435–451, Cham. Springer International Publishing.

Valls-Vargas, J., Zhu, J., and Ontañón, S. (2017). Graph grammar-based controllable generation of puzzles for a learning game about parallel programming. In *Proceedings of the FDG*, pages 1–10. ACM.

Walker, D. A. (2003). Jmasm9: Converting kendall's tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods Copyright*, 2:525–530.

Wouters, P., van Oostendorp, H., Boonekamp, R., and van der Spek, E. (2011). The role of game discourse analysis and curiosity in creating engaging and effective serious games by implementing a back story and foreshadowing. *Interacting with Computers*, 23(4):329 – 336.