It’s all a Skinner box? Typical Core Loop and Operant Conditioning

Thais Arrias Weiller
University of Tampere
Pontifícia Universidade Católica do Paraná
thais.arriasweiller@tuni.fi

ABSTRACT

This is an exploratory study on the hypothesis that a typical core loop for a videogame closely resembles an operant conditioning loop and, when that is not the case, the player’s experience will decline because of it. To assess this similarity, a double-blind experimental study was conducted in which participants were randomly assigned into two different groups and each group played a different version of the same decision-based game with 24 main choices. The experimental group (n = 55) received a version in which a classic core loop rewarded “correct” choices with mechanic and narrative rewards, while the control group (n = 54) experienced both mechanic and narrative punishments for the same “correct” choices. After the play session participants answered a questionnaire about their experience. The results showed a significant difference between groups’ experiences in certain dimensions, with an overall better experience being reported by the typical core loop group.

Keywords

Core loop, gameplay, game design, operant conditioning, mechanics, gameplay loop

INTRODUCTION

From the laurels of winning a match to the scorn of being checkmated, games poise a complex balance between humanity pleasure seeking behaviours and their tolerance for distress and difficulty (Bopp et al. 2016; Montola 2010). Although the intersection between both rewards and punishments of games have captured imagination and engagement for millennia, since the popularization of videogames this relationship has been much more scrutinized. Studies and theories present both positive and negative outlooks on the matter, ranging from videogames potential for learning (Batz et al. 2022; Gee 2008; Gee and Price 2021; Solanki and Mathew 2022) or evoking unique feelings to their players (Bopp et al. 2016; Bonk 2024; Iacovides and Cox 2015) but also their potential for harm (Vozaru 2020; Vu 2017). Despite their completely different hypothesis and methods analysing different segments of human experience and their connections to games, all these studies have in common the way they approach videogames: as if each one was a discrete and indivisible artefact. This presents a gap in the literature and, to better understand the relationship between games and their pleasures, we must analyse its elements and its effects on players.

© 2026 Authors & Digital Games Research Association DiGRA. Personal and educational classroom use of this paper is allowed, commercial use requires specific permission from the author.

This study takes an intersectional approach between two not commonly combined theorical frameworks to analyse the impact of a particular videogame element to the player experience. It understands games as complex systems composed by many different elements, many of which are particularly designed by its developers for an intended purposed (Sellers 2018). One of such elements, the *core loop*, is a game industry catch-all concept concerning the cycle of the most repeated actions throughout a game and its ensuing effects (Lemarchand 2021; Sellers 2018). The most common configuration for the core loop has a player's actions being followed by in-game consequences which, depending on the game design, can be understood as rewards or punishments (Katkoff 2013). This configuration has a distinguished similarity with a behaviour psychology concept of positive conditioning, in particular the positive reinforcement and punishment varieties (Fergusson 2010b; Murphy and Lupfer 2014). Operant conditioning loops are designed, in the same way a game is designed, to stimulate actions deemed desirable by designers while discouraging the ones deemed undesirable. This similarity has been pointed out previously in the literature (Fergusson 2010b; Katkoff 2013; Sellers 2018), but never fully investigated or correlated before.

To test this similarity, participants divided in two groups played different versions of the same decision-based game with 24 main choices, each group experiencing a version with dissimilar relationships between choices and their narrative and mechanical outcomes. The experimental group (n = 55) played a classic core loop with mechanical and narrative rewards for choices considered *correct* (method behind this definition is explained in the *Methods* section), whilst, in the version played by the control group (n = 54), the same correct choices resulted on negative mechanical and narrative outcomes. After playing their version of the game, participants answered a questionnaire about their experience (IJsselsteijn et al. 2013) using a 7-points Likert scale. These answers were then compared using a T-test to check how statistically significant their difference really was.

The contributions of this study are as follows: a) tests the hypothesis that a typical core loop and an operant conditioning loop are related as far as player experience is concerned, b) observes which dimensions of the player experience are impacted when the core loop does not resemble a conditioning loop and lastly c) finds which core loop arrangement provides a more pleasurable experience for the player.

OPERANT CONDITIONING

To better understand the relationship between this cole loops and that operant conditioning, it is first necessary to understand both of them apart. Succinctly, operant conditioning (or *instrumental* conditioning) postulates that voluntary behaviours can be stimulated or diminished depending on the consequences that follows them (Murphy and Lupfer 2014). The concept was shaped by American psychologist B.F. Skinner, who also created an apparatus to implement and test the theory on animals which was thus named *Skinner's box* (Honig and Staddon 2022). Operant conditioning differs from *classical* conditioning as the classical one targets *involuntary* behaviours, such as blinking (Kirsch et al. 2004). The use of both types of conditioning is frequently documented in the literature applied in situations in which a behaviour must be encouraged or extinguished, from involuntary ones such as a treatment for phobias or negative expression of psychiatric disorders (Kirsch et al. 2004), to voluntary such as to enhance learning (Rakasiwy et al. 2025).

There are different types of operant conditioning based on what type of consequences (*stimuli*) are associated with which behaviour, but most important to this research are the concepts of *positive reinforcement* and *positive punishment* (see Figure 1). Positive reinforcement is rewarding a desirable voluntary behaviour with a positive stimulus (Heidenreich 2007), which by its repetition, the individual comes to amalgamate the act of performing the behaviour with the pleasure derived from the reward. Positive punishment, on the other hand, is the same process of association with a stimulus, but using a punishment instead of reward, which results in a decrease of frequency in the behaviour.

What constitutes a reward or a punishment is quite ample and contextual. Food was used as reward for animals on many of Skinner's original tests, but the same food that was a successful reward for some can be a punishment for others (Honig and Staddon 2022). In humans in particular, a wide array of possible events or things can be considered a reward depending on the individual and the context. This is a very wide discussion, but in regard to this article subject of videogames, a reward is considered to be either something that positively impacts the player's capacity to continue to advance in the game (i.e. a status increase) or an outcome that the players values more than the alternative (i.e. a narrative resolution). In the same vein, a punishment in the context of this article is something that hinders player progress (i.e. status decrease) or presents outcomes that are less valued by the player (i.e. bad narrative occurrences).

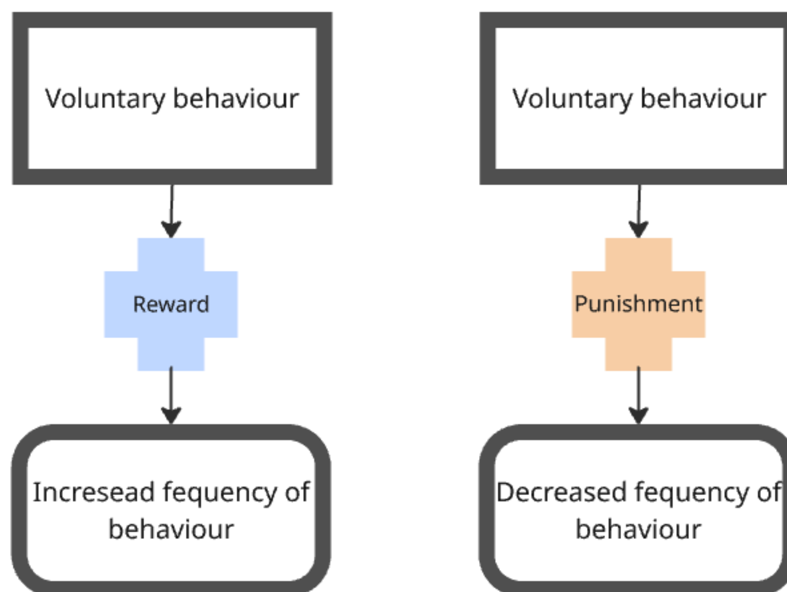

Figure 1: Diagrams for positive reinforcement and positive punishment.

If a reward is successfully interpreted as such and after enough recurrence, the behaviour itself becomes as pleasurable as the reward (Daw and Tobler 2014; Honig and Staddon 2022, 340), which leads to an increase frequency of said behaviour. An example of this process is when a music evokes feelings of joy by itself after being heard in repeated instances of joyful moments. In the same way, a punishment can also influence the sensation around the behaviour itself, as an example of the dread

some people experience when hearing their waking up alarm in unrelated situations. Both of those associations happen due to the individual's production of *expectancies*, as in the expectancy that the stimuli will happen again once the behaviour is repeated (Kirsch et al. 2004). Therefore, an action that was successfully conditioned will bring positive or negative sensations by itself.

The use of operant conditioning in videogames is not uncharted territory, from gambling (Kirsch et al. 2004) to addiction research (Vu 2017) there is a broad body of evidence already linking both fields. However, most of these studies approach the whole game as a stimulus (Järvelä et al. 2012; Vozaru 2020) or, when analysing in-game elements, resort to limited reward and punishment structures in game design features such as quests, peer recognition (Vu 2017) or reward schedules (Lorenz et al. 2015) without recognizing the underlying relationship of one of game design's most ubiquitous characteristics.

THE CORE LOOP

The core loop concept has been used in the industry at least since the early 2010 (Achterman 2011; Caldwell 2011; Nutt 2011; Staff 2011) and it became a staple of academic analysis (Ahmed and Noor 2024; ProzorR-Barbalat 2024) and education (Lemarchand 2021; Sellers 2018) since then.

The essence of a videogame rests in its interactivity, and when we isolate this interactivity on a single array of rules in which the player can act upon, we form what is called a mechanic (Järvinen 2007). Mechanics do not exist in a vacuum, they associate not only with the in-game reactions of the player's action but also with resulting status change in the game's presentation (Sellers 2018). These connections can happen by direct relationship (hitting an enemy causes it to lose health, or casting a fishing line can lead to catching a fish) or less direct ones (the total health of the enemy is lessened, or the inventory is now full of fish), but nevertheless these player actions and these game status changes are connected (Järvinen 2007). But these connections are not discrete; each status change invites a new player action, which could make use of the same mechanic or of another one, depending on this particular game's rules (Sellers 2018). Though sometimes the overall sum of these connections creates a linear progression, most commonly they concatenate in cycles, or loops, that can and will be repeated many times in a game.

These loops occur in many levels in a game, from broader to smaller instances (Kramarzewski and De Nucci 2023; Sellers 2018). The high-level loops represent larger, slower loops that take longer to finish (such as a whole game stage as a loop), while the low-level ones happen more often, with smaller time durations. The lowest level would be the cycle of actions the player performs at every given moment, with higher level loops being less repeated and other levels being ranked accordingly (Kramarzewski and De Nucci 2023; Sellers 2018). Therefore, a core loop would be a low-level loop, frequently repeated during the gameplay, making it an essential part of *all* interactive games.

As the core loop is the most repeated set of action in a game, designers devote extreme attention to its design (Katkoff 2013). Consequently, in most videogames, core loop actions that are positively valued by the game designer result in either a type of direct mechanic reward (extra points or extra items between other possible

alternatives) or a positive contextual reward (as a desirable narrative outcome or new narrative opportunities). On the other end of the spectrum, an action that is not positively valued by designers or deemed as the wrong action for the time or poorly executed is either ignored (such as in a match-3 game when the player tries to move an object to a position that does not lead to a match) or is actively punished.

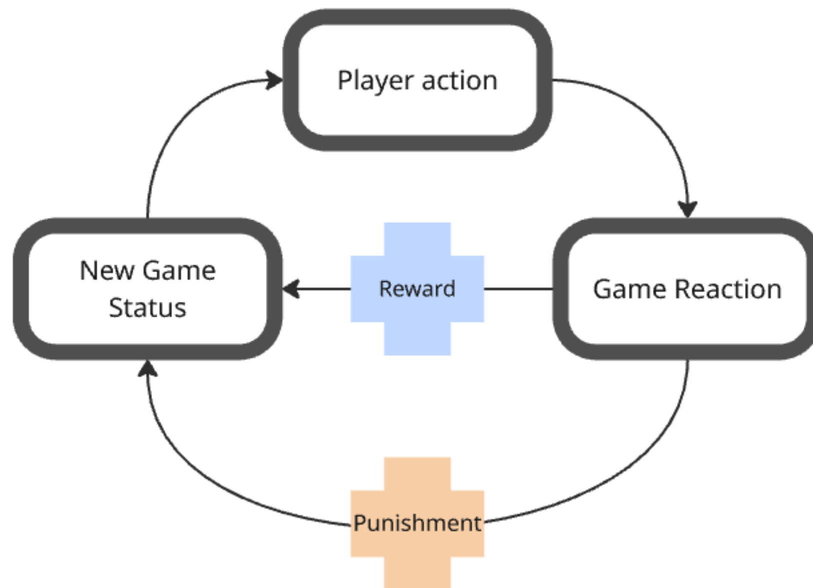

Figure 2: Representation of a core loop in which, depending on whether player's action is desirable or not, the game responds with either a reward or a punishment.

This construction (see Figure 2) sports a likeness to the positive reinforcement and punishment loops. This likeness was not ignored by the designers first discussing the core loop concept. Michael Fergusson wrote in two different developer texts about the intentional link between game loops and *compulsion loops*, which sport a striking resemblance with positive reinforcement loops. The first one was about the use of said loops to drive digital asset sales (Fergusson 2010a), and the second on the ethical implications of using such structures (Fergusson 2010b).

The benefits of having a core loop that resembles a conditioning loop are obvious. The central location of the core loop and its repeatability necessitate the player to interact with it more than any other mechanic loop in the game. If such a mediating part of the game was built in a way that positively conditions the player, this combination not only guides the player forward, but it also motivates their path in the designer's desired direction, while also disincentivizing unwanted behaviour and actions. Although the association between the core loop and conditioning was previously pointed out specially on industry contexts, such as in Fergusson's articles, no experimental data on that connection was found.

METHOD

To test the possible relationship between the two concepts, the exact same game needed to be used, differing solely by its reward and punishment parameters. The ability to produce this change for the player is not a common feature in most available commercial games. The next option considered was using a MOD[1] of an existing game, but concerns around if a MOD would allow to collect information about player's performance in the experience competed against this option. Finally, it was decided that, in order to have full control of the mechanics and access to all player data, the best experimental object would be to develop a new videogame, created just for this experiment, that was modelled after an existing game with a simple and discrete core loop that was already proven by its players as a successful one. The game *Reigns* (Nerial 2016) was chosen, as its gameplay met the requirements and presented a concise learning curve, perfect for a short experiment.

Reigns is a narrative driven game in which the player is presented by situations which have two options on how to proceed. The choice is made swiping left or right (akin to a dating app, such as Tinder), but, before releasing the finger, the player can have a preview of how each choice will impact on their reign's four statuses (represented by the population groups of clergy, peasantry, army and nobility). This preview, though not fully accurate, is important for player awareness, as reaching zero or maxing out in any of the four status leads to the player's demise. As a result, *Reigns* has a core loop that is composed by the actions of, firstly, reading and considering a situation and, secondly, picking between the two given choices. Such a straightforward core loop can disguise the complexity that emerges from the dynamics of playing it: each choice can impact more than one status, the extent of the influence is not completely predictable, and the player never knows which statuses the next situation will affect. Furthermore, choices are not made in a mechanical void: each presented situation is composed by a narrative context that colours the situation. These choices cannot be only seen as the unpassionate selection between stronger army and weaker church, but also the player's moral choice between repressing a community cult or raising church taxes and, moreover, the stoking of the tension between mechanical and narrative outcomes.

The Game

The game was made by the internal development team of the Pontifícia Universidade Católica do Paraná (PUCPR) according to author's specifications. The game replicated *Reigns* core loop of presenting a situation and giving the player two choices (see Figure 3 and 4). To make the experience more streamlined, the game had three statuses instead of *Reigns*' four. Reaching zero or maxing on any status meant game over.

In order to also test the narrative reward/punishment aspect, the game needed to bring some variety of moral decisions. Yet, the themes brought could not be too thematically charged in order not to emotionally vex any of the participants. It was decided that the best way to make the moral decision relevant to the participants without broaching taboo themes or making the choices too simplistic was to try to make them the most relatable possible for participants' lives. As the intended test site would be in a university, it needed to a theme that was pertinent to student's life but without alienating any other possible participants. The development team and the author decided then to use a travel theme, with the protagonist needing to get to a

friend's marriage (time pressure) together with two friends (social pressure) in an old car (material pressure). These three pressures became the themes for the narrative situations and their outcomes, while also becoming the three-status used for the mechanic dimension, zeroing or maxing out in any of them lead to game over. The game was made to be easily played on a laptop using the built in cursor due to its availability in the test site and its familiarity to the potential test audience.

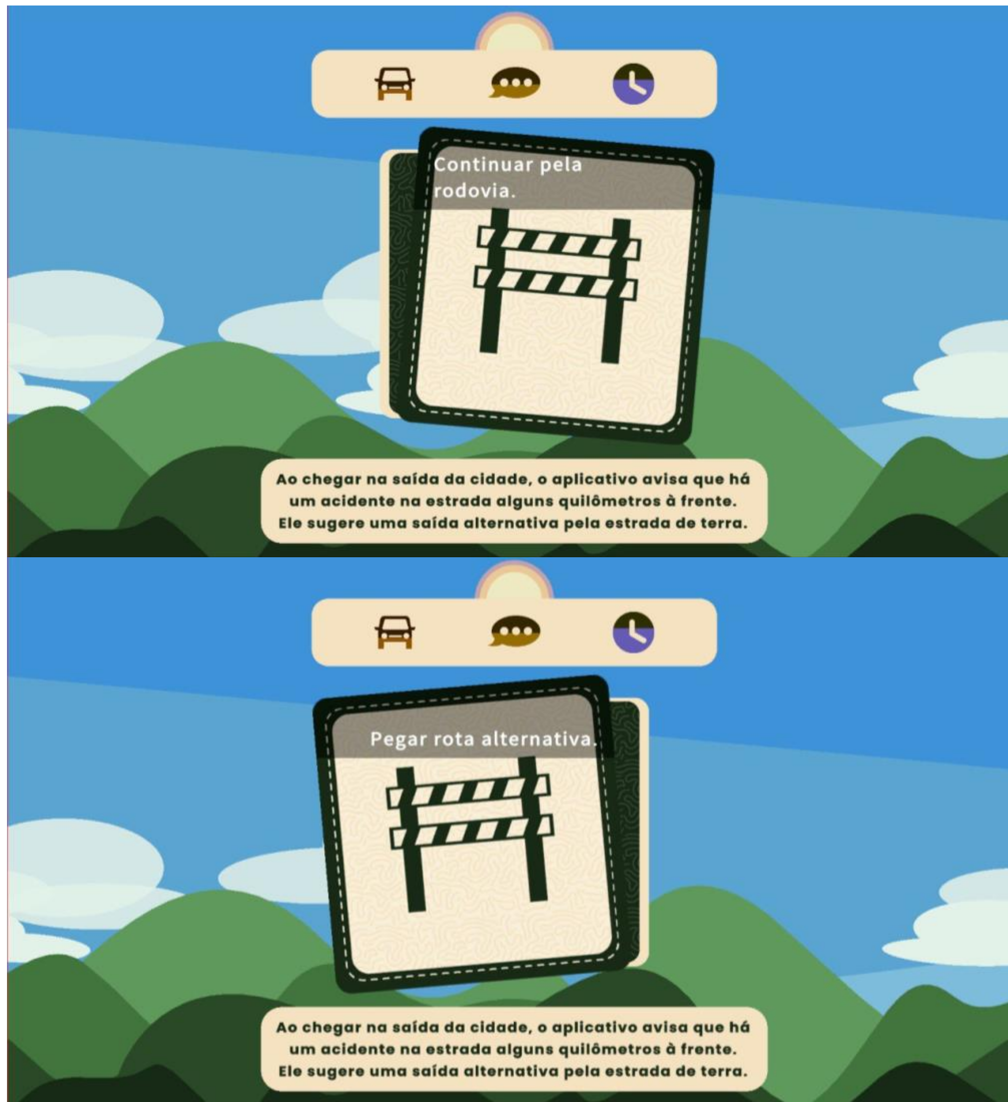

Figure 3 and 4: Participants used mouse movement to switch and visualize different options and mouse click anywhere on the screen to select the one currently displayed. The main situation text (at the bottom of both images) says "Upon reaching the city limits, the app warns of an accident on the road a few kilometres ahead. It suggests an alternative route via a dirt road". The choices (on the top of each figure, under the statuses) are "Continue through the asphalt road" for Fig. 3 and "Take the dirty road" on Fig. 4.

The main script had 24 main situations which had two possible options. Each of these choices led to a specific consequence card with its own text, which although also had two response options in order to keep reinforcing the same choice mechanic, both options lead to the same next main situation. Although all consequences were pre-determined by the choice made on a main situation, concerns about one single choice affecting two statuses at the same time being confusing to the player sometimes made the development team divide the mechanical feedback between the option choosing mechanic of the main situation and the consequence one. This allowed participants to better acclimate to the game, while also maintaining *both* mechanic and narrative consequences consistent to the main situation choice (see Figure 5 for reference of the flow).

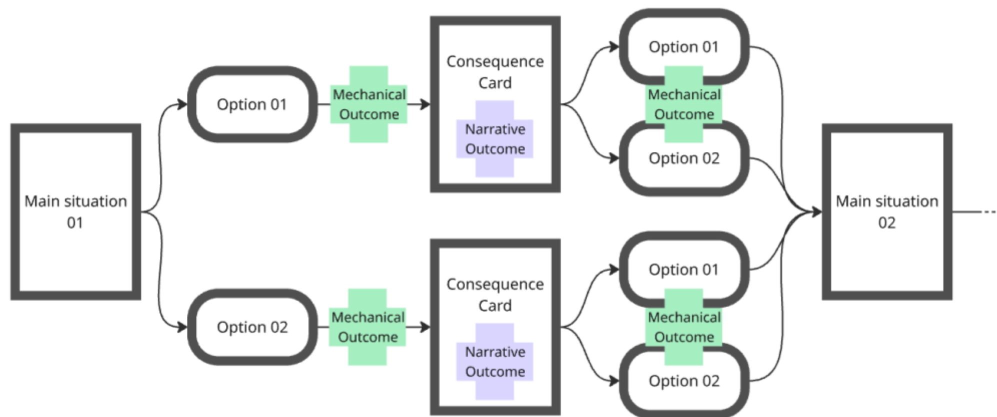

Figure 5: Player's decision flow and placement of the outcomes.

After the script for the main game was finalized and balanced as a typical core loop, work on the control version started. A total of four versions were made for this game, but only two are relevant for this article.

Versions

Each group played a version with dissimilar relationships between choices and mechanical and narrative rewards or punishments, one group playing a classic core loop with mechanic and narrative rewards for choices considered *correct*. On the other hand, in the version played by the other group, the same correct choices resulted on both mechanic and narrative punishment. This design was chosen because the removal of the outcome would make one of the versions significantly different from the other, experience wise, than the inversion of it (e.g. one version not having consequence texts).

Deciding which response could be considered more *correct* than the other is not a simple decision to make. Besides the many different moral philosophies that could be used to engage with this question (Miner and Petocz 2003), any one person's perspective could not align with the view of most of the participants of the study and thus producing skewed results or could reduce the choices to a naïve black-and-white worldview which could be considered preachy. Both concerns were heightened considering the study was conducted in a catholic university. To avoid presenting stale and simplistic moral questions or needing deep ethical discussion of which answer

was correct in the fictional scenario of the game, it was decided to rely on the model of intuitive morality and exemplars (MIME), which asserts an individual moral understanding is strongly influenced by their exposure to outside factors, such as media and social contexts (Tamborini 2012). To gauge the moral inclination of the future participants, all main situations were placed to the development team (n=10), many of which were students at PUCPR not long ago, with the phrasing "In your opinion, which is the most morally correct option?" for a consensus gauge. The team had as options the two choices available in the script plus a third one which indicated none of the options was more morally correct. From the 24 main situations, only 11 had the same answer being picked by 70% or more team members, and for that they were deemed our best educated guess of what a *correct* choice would be in the presented situation.

Those 11 situations in which one answer was deemed un-arguably more moral by the team were the answers that had their mechanic consequences switched and their narrative consequences cards rewritten in the control version while they were kept the same in the experimental one. They are presented on Table 1.

01 Unchanged	07 Unchanged	13 Unchanged	19 Unchanged
02 Unchanged	08 Changed	14 Changed	20 Unchanged
03 Changed	09 Changed	15 Unchanged	21 Changed
04 Changed	10 Changed	16 Unchanged	22 Unchanged
05 Changed	11 Unchanged	17 Changed	23 Unchanged
06 Unchanged	12 Changed	18 Changed	24 Unchanged

Table 1: Main situation with rewards and punishments switched on the control version. Typical core loop version kept all rewards and punishment as original.

The mechanical changes were straightforward: one choice that originally increased the social status two points decreased the same number of points in the control version. The narrative changes involved more work as both of the consequences cards needed to be re-written. For instance, on the Main Situation in which the player must choose between stopping to check what appears to be an accident to provide aid or to continue traveling without stopping, if they decide to stop on the normal core loop version, they find that the accident was not serious but the involved thank their support. Contrarily, on the control version, if the player decides to stop, they find there is no accident but only onlookers and their attempt was in vain. Most more moral options followed a similar pattern: the *correct* choice led not to a bad consequence, rather an unfulfilling and impotent one.

Participants

The participants (n = 109) were recruited from the public present in the Campus of PUCPR at the time of the experiment, which means they were mostly students and the age was young skewed. Group 01 (n = 55) ages ranged from 18 to 52 (mean = 23.5) and was composed by 42 participants which were students while group 02 (n = 54) ranged from 18 to 43 years old (mean = 22.7) and had also 42 student participants. Group 01 was composed by 25 men, 28 women and 2 non-binary people, while group 02 had 23, 29 and 2, respectively. After going through the experiment, participants were given a piece of candy as a thank you for their participation. Group 01 played

the normal core loop version (experimental) and group 02 played the one with outcomes switched (control).

Procedure

Before the beginning of the research, monitors were recruited from the Videogames Development degree from PUCPR. The monitor volunteers received credits for their work, and they were not briefed on the particularities of the study before the experiment to keep the study double-blind. After the experiment was finished, they not only were invited for a briefing about it but also received instruction on the nature of experimental research. The monitors were divided in two groups daily, one bigger group invited people present at the university's campus at the time to become participants and take part in the study, and the other smaller group, generally composed by one or two people, supervised the participants in the room in which the second and third stage of the experiment took place. Monitors from the second group also closed the questionnaire once the participant left and opened the game again, making the space ready for a new participant.

The whole experiment lasted five consecutive days, was conducted in the afternoons in three different rooms of the Yellow Building of PUCPR. The research had a total time between 15 and 25 minutes for the participant and happened in three stages.

First stage

Once a participant showed interest in taking part in the experiment, they were guided to the front of the room where the experiment was taking place on that day. Thus, the first stage began, and the participant was presented a simple overview of the entire experiment, followed by their informed consent[2] form and the collection of demographic data related to them, the latter two using an online form filled presently with the assistance of the main author. At this point, the participant was assigned a number that identified them for the remainder of the study.

Second stage

Participants were then led inside the room to play in the already placed laptops, and their first step starting the game was to enter their assigned number (see Figure 6). This number not only allowed us to record each player's journey playing but also defined which version of the game they played, while all monitors present at the time of the experiment ignorant about this information. This process maintained the double-blind during the experiment, as the author, the only person present aware of how the study worked, was outside the room while this and the next stage took place.

Participants were sat far enough from each other to make it hard to see other participants screens and, which could allow them to notice differences between their experiences. Monitors were instructed not to interfere with the participant's experience and only approach them if they noticed they were having difficulties or if they misbehaved. The number of simultaneous participants inside the room fluctuated depending on the number of available and willing people, but six participants played simultaneously at most.

Figure 6: First game screen, prompting participants for their number.

Third stage

When the participant ended the game or reached a game over, a screen thanking their contribution appeared and guided the participant to the third stage of the experiment (see Figure 7), which was the questionnaire to evaluate the players' experience. The questionnaire also only started after the participant identified their number.

Figure 7: End game screen, prompting participants to click on the button that took them to the questionnaire.

Data Collection Techniques

Besides the demographic information, collected only to add context to the experiment, there were two main methods of data collection in this study. The first one was the game itself, and the other was an experience questionnaire, answered right after playing the game.

Game Data

The game collected data about the play session and nested under the participant's number. For instance, it was collected time in which of play session started and each choice was made, the choice taken and resulting statuses and final screen before game over.

Game Experience Questionnaire

The test has a limited time, thus measuring repetition of the behaviour, as in animal tests, was not feasible. As the conditioning makes the behaviour itself become entangled with positive or negative emotions (Daw and Tobler 2014; Honig and Staddon 2022, 340), the participants' experience after the playing the videogame was used as measurement. The core model module of The Game Experience Questionnaire (IJsselsteijn et al. 2013) was used as a digital form to gauge participant's experience differences. All of its 33 sentences were followed by a 7-points Likert scale, presented as a scale from "Completely Disagree" to "Completely Agree". This module was chosen not only because it is highly accepted in the literature, but also for its sentences seemed more appropriated for this type of game. Other commonly used forms focus a lot on certain elements that are not relevant to this type of game, such as a gameplay induced time pressure (Denisova et al. 2020).

The GEQ recommends combining the results of each participant by 33 sentences on nested dimensions. The dimensions are:

1. Competence – The sense of prowess participants felt during the experience, associated with accomplishment and pride (Poels et al. 2009);
2. Sensory and imaginative immersion – The feel of being immersed in the game world (Poels et al. 2007);
3. Flow – Being concentrated and absorbed and losing contact with outside world (Poels et al. 2009);
4. Tension/annoyance – The feeling frustration or stress(IJsselsteijn et al. 2013);
5. Challenge – The sense of being challenged, of being on the edge of barely making (Poels et al. 2009);
6. Negative affect – Felling bored or disinterested (IJsselsteijn et al. 2013);
7. Positive affect – "probes the fun and enjoyment of gaming"(Poels et al. 2009, 21).

Though it was possible that *gameplay* element could be partially represented by the Challenge, Flow and Competence dimensions, it was decided it would be more direct to add six sentences to the original 33 to directly assess the main objective of this research. Four of those extra sentences were positive coded[3] and two were negative[4], which means they needed to be processed separately, as in positive ones a higher score means a better experience while on the negative ones the opposite is true.

Expected Results

It was expected that the control core loop group would demonstrate lower results on positive experiences related to gameplay (Competence, Flow and Positive Gameplay) and higher results on the negative ones (Challenge, Tension or Annoyance and Negative Gameplay). It was unknown if there would be any effect on the remaining

dimensions (Positive and Negative Affect and Imaginative Immersion) and they would probably be unchanged.

RESULTS

While processing the data collected, it was found that a significant portion of the participants (n = 51) met a game over condition before the 7th main situation, all of them due to maxing out in one of the statuses. Considering that the total number of 24 main situations, this meant these 21 participants from the experimental group and 30 from the control experienced less that 1/3 of the game. As having experienced so little of the game could result in an inaccurate sense of the game true experience, it was decided to proceed with only participants that progressed to the 7th situation or further, changing the total number of valid participants (n = 58). The 7th main situation was picked because it guaranteed that the participant had experienced at least 3 of the 11 main situations in which reward and punishment were switched, which meant they had had some contact with the experiment's main experimental variable.

Although there is some discussion whether Likert scales must be treated as ordinal or continuous data (Fagerland 2012), there seems to be consensus that if the data is an aggregation of rating scales, the distribution of the results approaches normality, the individual numerical scale is larger than 5 points and if the full scale of numerical values was used by participants, it is recommended to analyse it as continuous (Harpe 2015). As the data abided for all of these characteristics, parametric methods of analysis are believe to produce precise results (Fagerland 2012). In that understanding, Student's T-test is "relatively robust" for the required analysis (Harpe 2015, 842), and, accordingly, it was the test of choice to evaluate if the difference between dimension in each group was statistically significantly. As the sample was small and this is an exploratory study, a higher significance was chosen (<0.1).

Group Statistics				
	Version	Mean	Std. Deviation	Std. Error Mean
Competence	TCL (n = 34)	5.61	1.04	0.18
	CCL (n = 24)	5.04	1.31	0.27
Sensory and Imaginative Immersion	TCL (n = 34)	5.25	1.00	0.17
	CCL (n = 24)	5.13	1.23	0.25
Flow	TCL (n = 34)	4.62	1.28	0.22
	CCL (n = 24)	4.53	1.35	0.28
Tension/Annoyance	TCL (n = 34)	1.71	0.88	0.15
	CCL (n = 24)	1.94	1.10	0.22
Challenge	TCL (n = 34)	2.69	1.27	0.22
	CCL (n = 24)	3.15	1.06	0.22
Negative affect	TCL (n = 34)	2.18	0.95	0.16
	CCL (n = 24)	2.66	1.15	0.23
Positive affect	TCL (n = 34)	5.72	0.97	0.17
	CCL (n = 24)	5.63	1.22	0.25
Positive Gameplay	TCL (n = 34)	5.99	0.96	0.17
	CCL (n = 24)	6.04	0.99	0.20
Negative Gameplay	TCL (n = 34)	2.71	1.58	0.27
	CCL (n = 24)	3.38	1.64	0.34

Table 4: Descriptive of the data excluding participants which did not reach and/or surpass main situation 7. TCL stands for Typical Core Loop and CCL for Control Core Loop.

Independent Samples Test						
			Significance			
	F	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference
Competence	1.85	56.00	**0.04**	0.07	0.57	0.31
Sensory and Imaginative Immersion	0.44	56.00	0.33	0.66	0.13	0.29
Flow	0.26	56.00	0.40	0.80	0.09	0.35
Tension/Annoyance	-0.92	56.00	0.18	0.36	-0.24	0.26
Challenge	-1.45	56.00	**0.08**	0.15	-0.46	0.32
Negative affect	-1.74	56.00	**0.04**	0.09	-0.48	0.28
Positive affect	0.32	56.00	0.38	0.75	0.09	0.29
Positive Gameplay	-0.19	56.00	0.43	0.85	-0.05	0.26
Negative Gameplay	-1.56	56.00	**0.06**	0.12	-0.67	0.43

Table 5: T-test assuming equal variances of data excluding participants which did not reach and/or surpass main situation 7. Significant values in **bold.**

Conclusively, there is meaningful significance (<0.1) in the difference between the typical core loop and the control group in four dimensions. The dimension in which the significance is meaningful are Competence (0.03), Challenge (0.07), Negative Affect (0.04) and Negative Gameplay (0.06). The mean was lower on the control group when higher value meant better experience (5.61 vs 5.04 on Competence) and higher on the ones low numbers meant better experience (2.69 vs 3.15 on Challenge, 2.18 vs. 2.66 on Negative Affect and 2.71 vs 3.38 on Negative Gameplay) as theorized, and thus the use of the one-sided p. These results strongly indicate that participants on the control group felt less competent, more challenged and negative affected and experiencing unjoyful gameplay. Other dimensions showed little or no difference between groups.

We also looked for the relationship between dimensions across both groups (Table 6).

Inter-Item Correlation Matrix									
	Competence	Sensory and Imaginative Immersion	Flow	Tension/Annoyance	Challenge	Negative affect	Positive affect	Positive Gameplay	Negative Gameplay
Competence	1	0.35	0.37	-0.40	0.06	-0.30	0.57	0.33	-0.47
Sensory and Imaginative Immersion	0.35	1	0.69	-0.27	0.35	-0.42	0.78	0.57	-0.42
Flow	0.37	0.69	1	-0.24	0.29	-0.46	0.71	0.38	-0.31
Tension/Annoyance	-0.40	-0.27	-0.24	1	0.24	0.59	-0.44	-0.22	0.36
Challenge	0.06	0.35	0.29	0.24	1	0.27	0.30	0.06	0.00
Negative affect	-0.30	-0.42	-0.46	0.59	0.27	1	-0.51	-0.34	0.31
Positive affect	0.57	0.78	0.71	-0.44	0.30	-0.51	1	0.54	-0.53
Positive Gameplay	0.33	0.57	0.38	-0.22	0.06	-0.34	0.54	1	-0.44
Negative Gameplay	-0.47	-0.42	-0.31	0.36	0.00	0.31	-0.53	-0.44	1

Table 6: Relationship of dimensions between themselves.

Of remark is the strong relationship between Flow, Immersion and Positive Affect (all dual associations between these three dimensions are above 0.69). Though it could be said all these dimensions score positively and this relationship could be the sole result of that, Competence and Positive Gameplay also score positive and do not show such high association. No such relationship was found between the negative scoring dimensions, and the higher connection found (0.59) was between tension and Negative Affect.

DISCUSSION

This study intersectionally expanded upon our current understanding of game mechanics and its connection with other areas of expertise, such as psychology, informed by player experiences. We found a positive relationship between a core loop that is modelled after an operant conditioning loop to a better player experience, which indicates that this can be its typical format. Although the difference was not present in *all* dimensions analysed, this was expected as it was first theorized that not all of them were influenced by gameplay. The lack of change on those variables highlights how the difference in participants' experience was indeed only due to control changes.

The cognitive interpretation of the operant conditioning would attribute the lower experience rating given by the participants on the control group to the shattering of expectancies created by the switching of rewards and punishments. Once the participant believes one answer to be "correct" and the game punishes them for it, their enjoyment from the experience becomes irrevocably tarnished. This emphasizes how the conditioning format of the core loop functions as a type of *tacit contract* between designers and players, an unspoken bond in which each should perform their part, and the early established guidelines must be kept. Of course, the subversion of expectancies was also due to happen on the experimental group in cases in which the participants individual views and opinions differ from the ones present on the game, but as the election of the "correct" answers was made through a population pool, the number of main situations in which this happened must be much smaller.

	Changes in Control Group in Relation to the Experimental one	
	Theorized	Found
Competence	lower	significantly lower
Sensory and Imaginative Immersion	unknown	no relevant change
Flow	lower	no relevant change
Tension/Annoyance	higher	somewhat higher
Challenge	higher	significantly higher
Negative affect	higher	significantly higher
Positive affect	unknown	no relevant change
Positive Gameplay	unknown	no relevant change
Negative Gameplay	higher	significantly higher

Table 7: Outline of the theories and results found in the experiment.

Conclusively, according to these results, participants on the control group had a statistically significant worse experience in 4 of the 6 expected affected dimensions, a somewhat worse in one of them and presented no change in only one of the expected affected dimensions. Though Tension/annoyance did not present statically significant results, its numbers were not as similar in both groups as it was on the unchanged dimension, which means it could be more significant if a bigger pool of valid participants was used. Finally, from the three dimensions in which there was no expectation of change, no significant change was indeed found.

Although it is noteworthy that the control group – when controlled by players that did not experience enough with the game – was much smaller than the experiment group (10 participants less), which could skew the results, this difference can also be understood as yet another evidence of how the control core loop game made it considerably harder for participants to be able to progress.

Core Loop and Game Design research

Our results suggest that the closeness in which a core loop emulates a conditioning loop is linked with better player experience. This not only casts a new light on the understanding of player's experience and enjoyment in a game but also provides tools to further explore these connections. As a result, future research on core loops can use conditioning loops as an understudy for the core loop when player experience is relevant, as long as it is accounted for the dimensions in which difference was not significant.

Game Design Research, Psychology and Ethics

Although there is a rich tapestry of studies on the relationship of games in general and videogames in particular with psychological effects (Abuhamdeh et al. 2015; Bopp et al. 2016; Daw and Tobler 2014; Weiller et al. 2024) and conditioning (Lewis et al. 2012; Rakasiwy et al. 2025; Sault 2023; Vu 2017) or the possibility of the whole game being a stimulus for an outside behaviour (Järvelä et al. 2012; Vozaru 2020), there is still very few studies tackling the role of a game's design components in connection with the player's experience. Finding the relationship of a particular design element, the core loop, with operant conditioning allows us to analyse in a more precise manner the link of this particular element of games with specific reactions in players, which furthers our understanding of both psychology and game design. Since our results indicate that game loops can be understood as conditioning loops, they entail that a new multitude of psychological tests currently used in behaviourist literature could also be used in investigating particular game elements. For instance, experimental methods such as shaping, schedules of reinforcement, extinction, behavioural chains or stimulus control to be used to further our understanding of game loops or even other elements such gameplay, level design or difficulty. Likewise, other elements in game design can have links to other psychological concepts.

These results also underscore the ethical implications of making games. If all games have core loops and core loops can be operant conditioning loops, it is within reason that all games have the potential to condition the player to continue playing them.

Limitations and Future Work

Although we collected a sample size that would suffice for a traditional t-test, the necessity of accounting for the participants which did not progress enough in the game resulted in a much smaller sample size (n=58). As a result, even though the typical core loop group added up to a sufficiently sized sample for this methodology (n = 34), the control group (n = 24) ended up numbered under the recommendation of 30 samples. While this could have skewed the data, as the original data without the removals showed the same tendencies (in a sample size >50).

While the results are significant, it needs to be pointed out that only a story driven core loop was tested in this experiment. Although there is no reason to believe other types of games would have significantly different results, it still would be interesting to expand this type of research to different genres of games and different applications od core loops. The cultural contexts and social economic make-up of the participants could also have a big role in these results. Similar tests should be undertaken in other countries and under different situations that are not a university to control for that influence.

As the early conclusion of the game had an attenuating effect on participant's experience and was the reason why so many participants did not experience the tested variables and had to be disregarded of the final tabulation as a result, we want to repeat this experiment with a different game balancing, that would allow for longer gameplay. Although these results are encouraging, a bigger pool of valid participants would also allow future results to be even more conclusive.

CONCLUSION

This article proposed to illuminate the relationship between the core loop and an operant conditioning loop using player experience as validation. To do so, the same game with different relationships between mechanic and narrative punishments and rewards was played by participants (n = 109) in order to find if there is significant difference of experience between players of each game version. The data collected, when controlled for participants that did not experience enough of the game (n = 58), shows a strong relationship between a core loop which has a positive reinforcement loop composition and a better player experience in most player experience dimensions related to gameplay. These results not only add to our understanding of game elements and design but also can serve as foundation for new research avenues on the complex relationship between games and player's pleasure and annoyance.

ACKNOWLEDGMENTS

This study was approved by PUCPR Ethics Committee under protocol number 91436125.8.0000.0020. We thank all the participants for their time, but specially we thank the monitors that assisted the experiment and the developers behind the game. The game developing team was composed by Joelson Barbosa Ferreira, Tiago Jonathan de Deus Souza and Giovana Marchiori Ribeiro, with additional help from Felipe Cortizo Elias e Silva, Henrique Resende Specian, Anderson Cardoso and Maurício Perin.

Monitors were Arthur Stumpf, Beatriz Oliveira dos Santos, Bruno Henrique Kulik Barbosa, Carla Rafaela Fernandes Cardoso, Carlos Eduardo Petrykoski Bueno, Carolina Fernandes de Deus, Diego Matheus dos Santos, Enzo Bonaccordi Lissoni, Gabriel Pelisson, Giovana Mocelin Costa, Heber Levantino de Souza Junior, Igor de Jesus Soares Fernandes, Jonatas da Costa Nascimento, Lara Foggiatto Andriguetto, Leonardo Ferreira Rocha, Leonie Figueiredo Morona, Lucas Eduardo de Camargo Heisler, Mariane de Lima Mesquita, Massao Oshima, Paulo Enrique Cassinelli, Pedro Henrique Schneider Fagundes de Oliveira, Rafael Guilherme Krueger, Rafael mielke schwartsburd, Ricardo da Rocha Carvalho, Rodrigo Ribas Godoy, Vinicius Martello da Silva and Vittório Bucci Conte.

REFERENCES

Abuhamdeh, Sami, Mihaly Csikszentmihalyi, and Baland Jalal. 2015. 'Enjoying the Possibility of Defeat: Outcome Uncertainty, Suspense, and Intrinsic Motivation'. *Motivation and Emotion* 39 (1): 1–10. https://doi.org/10.1007/s11031-014-9425-2.

Achterman, Daniel. 2011. 'The Craft of Game Systems: General Guidelines'. Game Developer, November 18. https://www.gamedeveloper.com/design/the-craft-of-game-systems-general-guidelines.

Ahmed, Shamit, and Faiza Anan Noor. 2024. 'Rapid Game World Creation Leveraging Geospatial Data in Game Jams: The 'Galactic Gourmet' Post-Mortem'. *Proceedings of the 8th International Conference on Game Jams, Hackathons and Game Creation Events*, October 11, 76–80. https://doi.org/10.1145/3697789.3697804.

Batz, Victoria, Inga Lipowski, Matthias Morfeld, Christian Hansen, and Michael A. Herzog. 2022. 'Accessible Design of Serious Games for People with Intellectual Disabilities in Inclusive Vocational Education'. Paper presented at Proceedings of DiGRA 2022 Conference: Bringing Worlds Together. January 1. https://doi.org/10.26503/dl.v2022i1.1372.

Bonk, Devin Nicholas. 2024. 'Emotions, Moods, and Videogames: Exploring the Affective Experience of Gamers'. Faculty of Kinesiology and Physical Education University of Toronto.

Bopp, Julia Ayumi, Elisa D. Mekler, and Klaus Opwis. 2016. 'Negative Emotion, Positive Experience?: Emotionally Moving Moments in Digital Games'. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, May 7, 2996–3006. https://doi.org/10.1145/2858036.2858227.

Caldwell, Brendan. 2011. 'The Reasons Behind SpyParty'. January 10. https://www.gamedeveloper.com/design/the-reasons-behind-i-spyparty-i-.

Daw, Nathaniel D., and Philippe N. Tobler. 2014. 'Value Learning through Reinforcement'. In *Neuroeconomics*. Elsevier. https://doi.org/10.1016/B978-0-12-416008-8.00015-2.

Denisova, Alena, Paul Cairns, Christian Guckelsberger, and David Zendle. 2020. 'Measuring Perceived Challenge in Digital Games: Development & Validation of the Challenge Originating from Recent Gameplay Interaction Scale (CORGIS)'. *International Journal of Human-Computer Studies* 137 (May): 102383. https://doi.org/10.1016/j.ijhcs.2019.102383.

Fagerland, Morten W. 2012. 'T-Tests, Non-Parametric Tests, and Large Studies—a Paradox of Statistical Practice?' *BMC Medical Research Methodology* 12 (1): 78. https://doi.org/10.1186/1471-2288-12-78.

Fergusson, Michael. 2010a. 'Achievements, Social Games and Virtual Goods'. Game Developer, September 7. https://www.gamedeveloper.com/design/achievements-social-games-and-virtual-goods.

Fergusson, Michael. 2010b. 'Playing with Fire: Ethics and Game Design'. December 1. https://www.gamedeveloper.com/design/playing-with-fire-ethics-and-game-design.

Franceschini, Sandro, Sara Bertoni, Matteo Lulli, Telmo Pievani, and Andrea Facoetti. 2022. 'Short-Term Effects of Video-Games on Cognitive Enhancement: The Role of Positive Emotions'. *Journal of Cognitive Enhancement* 6 (1): 29–46. https://doi.org/10.1007/s41465-021-00220-9.

Gee, James Paul. 2008. *What Video Games Have to Teach Us about Learning and Literacy*. Rev. and Updatet ed. Palgrave Macmillan.

Gee, James Paul, and Amy Price. 2021. 'Game-Design Teaching and Learning'. *Strategies* 34 (3): 35–38. https://doi.org/10.1080/08924562.2021.1896928.

Harpe, Spencer E. 2015. 'How to Analyze Likert and Other Rating Scale Data'. *Currents in Pharmacy Teaching and Learning* 7 (6): 836–50. https://doi.org/10.1016/j.cptl.2015.08.001.

Heidenreich, Barbara. 2007. 'An Introduction to Positive Reinforcement Training and Its Benefits'. *Journal of Exotic Pet Medicine* 16 (1): 19–23. https://doi.org/10.1053/j.jepm.2006.11.005.

Honig, Werner K., and J. E. R. Staddon. 2022. *Handbook of Operant Behavior*. 1st edn. Routledge. https://doi.org/10.4324/9781003256670.

Iacovides, Ioanna, and Anna L. Cox. 2015. 'Moving Beyond Fun: Evaluating Serious Experience in Digital Games'. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, April 18, 2245–54. https://doi.org/10.1145/2702123.2702204.

IJsselsteijn, W. A., Y. A. de Kort, and Y. Poels. 2013. 'The Game Experience Questionnaire'. Technische Universiteit Eindhoven.

Järvelä, Simo, Inger Ekman, J. Matias Kivikangas, and Niklas Ravaja. 2012. 'Digital Games as Experiment Stimulus'. Paper presented at Proceedings of Nordic DiGRA 2012 Conference. *Proceedings of Nordic DiGRA 2012 Conference*, January 1. https://doi.org/10.26503/dl.v2012i1.607.

Järvinen, Aki. 2007. 'Introducing Applied Ludology: Hands-on Methods for Game Studies'. *Situated Play, Proceedings of DiGRA 2007 Conference*, 134–44.

Juul, Jesper. 2003. 'The Game, the Player, the World: Looking for a Heart of Gameness'. *DiGRA '03 - Proceedings of the 2003 DiGRA International Conference: Level Up* 2 (1): 30–45.

Katkoff, Michail. 2013. 'Mid-Core Success Part 1: Core Loops'. October 24. https://www.gamedeveloper.com/design/mid-core-success-part-1-core-loops.

Kirsch, Irving, Steven Jay Lynn, Michael Vigorito, and Ralph R. Miller. 2004. ‘The Role of Cognition in Classical and Operant Conditioning’. *Journal of Clinical Psychology* 60 (4): 369–92. https://doi.org/10.1002/jclp.10251.

Kramarzewski, Adam, and Ennio De Nucci. 2023. *Practical Game Design: A Modern and Comprehensive Guide to Video Game Design*. Second edition. Packt Publishing Ltd.

Kuittinen, Jussi, and Jussi Holopainen. 2009. ‘Some Notes on the Nature of Game Design’. Paper presented at Breaking New Ground: Innovation in Games, Play, Practice and Theory. *Proceedings of DiGRA 2009*.

Lemarchand, Richard. 2021. *A Playful Production Process: For Game Designers (and Everyone)*. The MIT Press.

Lewis, Chris, Noah Wardrip-Fruin, and Jim Whitehead. 2012. ‘Motivational Game Design Patterns of ’ville Games’. *Proceedings of the International Conference on the Foundations of Digital Games*, May 29, 172–79. https://doi.org/10.1145/2282338.2282373.

Lorenz, Robert C., Tobias Gleich, Jürgen Gallinat, and Simone Kühn. 2015. ‘Video Game Training and the Reward System’. *Frontiers in Human Neuroscience* 9 (February). https://doi.org/10.3389/fnhum.2015.00040.

Miner, Maureen, and Agnes Petocz. 2003. ‘Moral Theory in Ethical Decision Making: Problems, Clarifications and Recommendations from a Psychological Perspective’. *Journal of Business Ethics*, no. 42.

Montola, Markus. 2010. ‘The Positive Negative Experience in Extreme Role-Playing’. *Nordic DiGRA*.

Murphy, Eric S., and Gwen J. Lupfer. 2014. ‘Basic Principles of Operant Conditioning’. In *The Wiley Blackwell Handbook of Operant and Classical Conditioning*, 1st edn, edited by Frances K. McSweeney and Eric S. Murphy. Wiley. https://doi.org/10.1002/9781118468135.ch8.

Nerial. 2016. *Reigns*. Mobile Game. London, UK: Devolver Digital.

Nutt, Christian. 2011. ‘Infinity Blade: A New Era Of Games’. April 11. https://www.gamedeveloper.com/design/-i-infinity-blade-i-a-new-era-of-games.

Poels, K., Y. A. W. de Kort, and W. A. IJsselsteijn. 2009. *Deliverable D3.3: GAME EXPERIENCE QUESTIONNAIRE*. No. 028765. TU/e.

Poels, Karolien, Yvonne De Kort, and Wijnand Ijsselsteijn. 2007. ‘“It Is Always a Lot of Fun!”: Exploring Dimensions of Digital Game Experience Using Focus Group Methodology’. *Proceedings of the 2007 Conference on Future Play - Future Play ’07*, 83. https://doi.org/10.1145/1328202.1328218.

ProzorR-Barbalat, Liliana. 2024. ‘Common Mistakes Game Developers Make’. (Chișinău, Republica Moldova), 2264–66.

Rakasiwy, Mochamad Randy Pradika, Viona Dea Surya Cipta, and Nanditas Adhys Nur Fadhilah. 2025. ‘The Role Of Operant Conditioning In Enhaching Learning Outcomes In Social Studies Education’. *SOSEARCH : Social Science Educational Research* 6 (1): 38–49. https://doi.org/10.26740/sosearch.v6n1.p38-49.

Sault, DM. 2023. *The Influence of Videogame Reinforcement Schedules on Game Play Duration*. 773499 Bytes. 773499 Bytes. https://doi.org/10.25959/23240732.V1.

Sellers, Michael. 2018. *Advanced Game Design: A Systems Approach*. Addison-Wesley.

Solanki, Chaitanya, and Deepak J. Mathew. 2022. 'Habit and Perception towards Games and Game-Based Learning in India: A Survey Study of Awareness and Experience with Digital Games'. Paper presented at Proceedings of DiGRA 2022 Conference: Bringing Worlds Together. January 1. https://doi.org/10.26503/dl.v2022i1.1337.

Staff. 2011. 'Behind The Scenes With People Can Fly's Bulletstorm'. October 11. https://www.gamedeveloper.com/design/behind-the-scenes-with-people-can-fly-s-i-bulletstorm-i-.

Tamborini, Ron, ed. 2012. *Media and the Moral Mind*. 0 edn. Routledge. https://doi.org/10.4324/9780203127070.

Vozaru, Miruna. 2020. 'From Stimulus to Situation – Tracking Situational Information Manipulation in Quantitative Intervention-Based Game Research'. *Abstract Proceedings of DiGRA 2020 Conference: Play Everywhere*.

Vu, Daniel. 2017. 'An Analysis of Operant Conditioning and Its Relationship with Video Game Addiction'. *ART 108: Introduction to Games Studies*, December 1. https://scholarworks.sjsu.edu/art108/6.

Weiller, Thais Arrias, Pedro Campos, and Vanessa Cesário. 2024. 'The Role of Negative Emotions in Videogames'. Paper presented at Human Interaction and Emerging Technologies (IHIET 2024). https://doi.org/10.54941/ahfe1005474.

ENDNOTES

[1] MOD is an alteration to an original game made by a third party, generally a fan, in order to change an aspect of a game. Though many MODs are aesthetic (i.e. exchanging original graphics for higher textured ones or exchanging for something else entirely), many are functional, and some allow players to even use an original game as a tool to create different experiences.

[2] The informed consent here concerns the definition used by PUCPR Ethics Committee, which requires researchers to inform participants about the experience, total time duration of test, themes and activities concerned, risks and benefits, anonymity procedures, possible triggering information, contact information for the researchers and for the committee itself and what type of research the test is about. It does not require to inform in detail the test design, so participants, like the monitors, did not know that there were different versions of the game and how these versions were different.

[3] The sentences were: "I like games like this", "The core mechanic was simple to understand", "It was easy to understand what to do in the game" and "I'd like to play this game again".

[4] The sentences were: "The consequences felt random" and "I got confused about what the game expected of me".