



**HAL**  
open science

## **Difficulty influence on motivation over time in video games using survival analysis**

Thibault Allart, Guillaume Levieux, Michel Pierfitte, Agathe Guilloux, Stéphane Natkin

### ► **To cite this version:**

Thibault Allart, Guillaume Levieux, Michel Pierfitte, Agathe Guilloux, Stéphane Natkin. Difficulty influence on motivation over time in video games using survival analysis. FDG - Foundation of Digital Games, Aug 2017, Hyannis, United States. <10.1145/3102071.3102085>. <hal-02436676>

**HAL Id: hal-02436676**

**<https://hal.science/hal-02436676v1>**

Submitted on 13 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Difficulty Influence on Motivation over Time in Video Games using Survival Analysis

Thibault Allart  
Ubisoft  
126 rue de Lagny  
Montreuil 93100, France  
thibault.allart@gmail.com

Guillaume Levieux  
CNAM / CEDRIC  
292 rue Saint Martin  
Paris 75003, France  
guillaume.levieux@lecnam.net

Michel Pierfitte  
Ubisoft  
126 rue de Lagny  
Montreuil 93100, France  
michel.pierfitte@ubisoft.com

Agathe Guilloux  
UEVE-PSaclay/LaMME  
23 Boulevard de France  
Evry 91037, France  
agathe.guilloux@math.cnrs.fr

Stephane Natkin  
CNAM / CEDRIC  
292 rue Saint Martin  
Paris 75003, France  
stephane.natkin@lecnam.net

## ABSTRACT

In this paper, we study the link between difficulty and player's motivation in two games developed by Ubisoft®: *Rayman® Legends* and *Tom Clancy's The Division®*. We describe a method to estimate players' difficulty over time and link it's time varying effect with players retention. Results confirm flow and self-efficacy theory. Also, for the first hours of playtime, results differ between the two games. We explain that discrepancy with regard to attribution theory : in *Rayman Legends*, failure can be mainly attributed to the player skills, while in *Tom Clancy's The Division*, avatar's strength plays a fundamental role and can always be relatively quickly improved.

## CCS CONCEPTS

•Mathematics of computing → Survival analysis; •Applied computing → Computer games;

## KEYWORDS

Game Analytics, Difficulty, Motivation, Survival Analysis, Video Games

### ACM Reference format:

Thibault Allart, Guillaume Levieux, Michel Pierfitte, Agathe Guilloux, and Stephane Natkin. 2017. Difficulty Influence on Motivation over Time in Video Games using Survival Analysis. In *Proceedings of FDG'17, Hyannis, MA, USA, August 14-17, 2017*, 6 pages.  
DOI: 10.1145/3102071.3102085

## 1 INTRODUCTION

When it comes to foster and maintain a video game player's motivation, difficulty is one of the most important parameter to carefully adjust. Authors have tried to explain the inherent appeal of video games, and many of them consider challenge as one of its most fundamental characteristics. However, challenge or difficulty are

complex subjective notions, and their relationship with motivation is not obvious. In this paper, our aim is to investigate this relationship, and thus help the design of more compelling interactive experiences.

To study the motivational aspects of video games, many authors observe a recruited sample of players during a controlled experiment, sometimes even using games specifically designed for the study. This type of study allows A/B testing and fine-grained observation of the players' behavior. However, such an experimental setting has a certain cost, and cannot be maintained for a long period of time.

However, when it comes to challenge and difficulty, time seems to be an important parameter. Of course, the player's knowledge and skills are changing over time and thus the absolute difficulty needs to rise with them. But even from a relative point of view, the players may not appreciate the same level of relative difficulty at the beginning of the game, when they totally discover the gameplay mechanics and at the end, when they combine the basic mechanics to discover more advanced gameplay dynamics. The problem is that, from an experimental point of view, the impact of difficulty over time on motivation might be very hard to observe in a controlled experiment, as the players' progression may require many hours of playtime.

Hopefully, out of laboratories, player monitoring is becoming standard. Gaming hardwares are today almost always connected to the Internet. Game sessions can be recorded, and gathered by game companies. These data allow game developers to get a more detailed feedback on their work, and even to update their game after release, according to the players behavior. These recordings allow us to study players that have not been recruited, and that play polished gameplays for long period of times.

This gives us a chance to see how motivational models behave in a real-world gaming context, over large periods of time.

In the first sections of the paper, we describe how we compute difficulty for both games, while explaining in detail the survival analysis model we developed for this study. Then, we analyze the link between difficulty and player's retention measured in both games. Finally, we explain how these results may confirm motivational theories in the context of video games.

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.  
*FDG'17, Hyannis, MA, USA*

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
978-1-4503-5319-9/17/08...\$15.00  
DOI: 10.1145/3102071.3102085

## 2 DIFFICULTY IN VIDEO GAMES

Many authors consider challenge as one of the most fundamental aspect of video games' inherent appeal. Malone has studied young children's motivation to play video games. He proposes three features that make computer games so captivating: challenge, curiosity and fantasy [15]. In his model, challenge is directly related to the game's difficulty and corresponds to the uncertainty for the player to reach the game's goals. Lazzaro proposes a four factor model, where *Hard Fun* is related to the feeling of overcoming difficult tasks [13]. Sweetser et al see challenge as one of the most important part of their Game Flow framework [18]. The work of Sweetser et al stems from Mihaly Csikszentmihalyi's Theory of Flow [7]. Csikszentmihalyi has been trying to figure out the properties of activities showing a strong, intrinsic ability to motivate. His research states that these activities provide *perceived challenges, or opportunities for action, that stretch (neither overmatching nor underusing) existing skills* [7]. Ryan et al also studied intrinsic motivation and apply their Self-Determination Theory to video games, showing how enjoyment is related to the feeling of *competence*, which relies on an optimal level of challenge, and thus, to the game's difficulty [17]. Jesper Juul provided insight on how failure, and thus difficulty, is one of the core aspects of video game enjoyment and learning progression [10].

However, while everyone seems to agree on the fact that difficulty is a central aspect of video games, it is still very hard to accurately describe the relationship between difficulty and motivation in video games. First, because difficulty, as well as motivation, are complex notions that can't be directly measured. We can use questionnaires to assess the subjective aspects of difficulty and motivation but each measure will interrupt the game and thus this assessment cannot be made very often. This is even more problematic if we want to assess the impact of difficulty on motivation over time, because we hardly can rely on a post experiment questionnaire to accurately evaluate the evolution of difficulty throughout the game. It is possible to integrate such an assessment to the gameplay, as Constant et al did for the subjective difficulty [5] But this kind of assessment is restricted to turn based games and necessitates a gameplay modification.

Then, we can try to estimate difficulty and motivation from the player's actual behavior in the game. First, the player's performance is related to the game's difficulty. If the game is harder, the player will lose more often, and this failure ratio can be monitored. However, as detailed by Levieux et al, this measure of difficulty is also linked to the player's motivation [14]. Indeed, the player's performance can decrease both because the game is more demanding or because the player's motivation has dropped, and he invests less attentional resources in the game. We thus have to make rough assumptions to study a game's difficulty using the player's performance: that while he is playing, he always does his best to win the game, and that if his motivation drops below a certain level, he will just stop playing. With this assumption, we can consider that the player's performance allows us to estimate the objective difficulty of the game, and that the link between difficulty and motivation is similar to the link between difficulty and retention.

Many research point the importance to understand the temporal aspect of the game [12] and propose a way to model it [21]. In

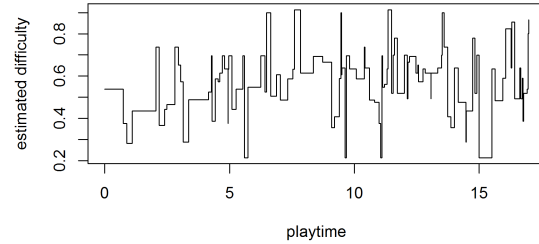


Figure 1: Estimated difficulty of one *Rayman Legends* player

this paper, we are particularly interested in the impact over time of difficulty on motivation. First, it is to note that difficulty varies with time in a video game. The player gets better at playing, but the game also constantly increases the required performance level. Moreover, from a motivational perspective, difficulty has not the same impact on motivation at the beginning or at the end of the game. Klimmt et al realized an experiment on a shooter game where player enjoyed the lowest level of difficulty. They suggest that at the beginning of a game, the players' representation of the game's difficulty is not calibrated, and that they tend to like high amounts of positive feedback and do not realize that difficulty is low. Having a low difficulty at the start of the game is also coherent with self-efficacy theory: Albert Bandura explains that failure lowers mastery expectations, especially in the early course of events. If failure happens after a strong efficacy expectation has been developed, then failure should have a smaller negative impact on motivation, or even have a positive one if the players know that failure can be overcome with sustained effort [2].

The influence of difficulty over time can be hard to capture in a controlled experiment, primarily because of the complexity and cost of following players over long period of times. Hopefully, video games remote monitoring gives us access to the whole history of players' behavior. In this paper, we describe a method to analyze such longitudinal data, and apply this method to two video games where players were remotely monitored during their game sessions.

### 2.1 Estimating difficulty

We estimate the difficulty of a challenge from the players' failures and successes for that challenge, as Levieux et al defined it [14]. Let  $X$  being game variables that represent both player and challenges characteristics, then we estimate the difficulty as the probability to fail knowing those characteristics.

$$\mathbb{P}(\text{fail}|X)$$

At this point, any statistical model can be used to model difficulty. However, to model the difficulty of a challenge, we need variables describing both the challenge and the player skills. Telemetric data often describe challenge completion, but are rarely accurate enough to describe the player's mastery of the game mechanics, i.e., the player skills. Thus, we need to take into account skills variability among players, that is, the fact that a challenge will not have the same difficulty for each player. We make the assumption that

skills variability is randomly distributed among players, following a normal distribution. Consequently, we can model a challenge difficulty and take into account this unobserved random effect by using a mixed effect logistic regression. Challenge characteristics are modeled as fixed effects and a random intercept is to take into each player's skills. Once we get a model of player's difficulty associated to a given challenge, player's characteristics and player's skill, we can compute the estimated difficulty.

Our difficulty model allows us to estimate the game's difficulty over time. Indeed, for each value of playtime, we know the value of game variables associated to the current challenge. Thus, we can compute the estimated difficulty over playtime<sup>1</sup> for each player as illustrated on figure 1. A difficulty of 0.3 means that player probability to fail the challenge knowing his characteristics is of 30%.

### 3 SURVIVAL ANALYSIS

#### 3.1 Retention

Motivation is a complex psychological construct that we cannot directly observe. In this research, we estimate players' motivation through retention. Indeed, we suppose that the lowest the motivation, the more chances there are the player stops playing.

Many authors have studied the impact of design variables on retention [1, 3, 8, 9, 11, 22]. None of this research has however focused on the impact of difficulty on retention. Our aim here is not to predict retention but rather to be able to better describe the link between difficulty and retention, and thus, player's motivation.

To do so, we rely on survival analysis. For an overview of survival analysis applied to playtime measurement see [20]. Let  $T$  be a random variable associated to player's total playtime. Retention is defined as the amount of players remaining in the game after  $t$  hours of playtime and is noted  $S(t)$ .

$$S(t) = \mathbf{P}(T > t) \quad (1)$$

In our research, we estimate the link between difficulty and retention using a regression model that relies on a slightly different way to look at retention: the quit rate. Quit rate, noted  $\lambda(t)$ , is the probability that a player stops playing and never plays again afterwards, in a small interval of time, knowing that he was playing just before. Among many authors, it was used by Chen et al to model the impact of network quality on player departure [4].

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{\mathbf{P}(t < T \leq t + h | T > t)}{h} \quad (2)$$

It is well know that we can compute quit rate from retention and vice-versa. A natural way to model the influence of game variables on quit rate is to use the Cox proportional hazard model [6].

Equation (3) shows an extend version of Cox proportional hazard that includes time varying variables and time varying coefficients.

$$\lambda(t) = \lambda_0(t) \exp \left( \sum_{j=1}^p X_{i,j}(t) \beta_j(t) \right) \quad (3)$$

Various approaches have been proposed to deal with time varying variables [24], time varying coefficients [25] or both [16]. However, none of them include both time varying variables and time varying coefficients and can also be scaled up to deal with the very large amount of data available in the video game industry.

#### 3.2 Cox model parameter estimation

As for many statistical methods, we search for parameters values that maximize the probability to have seen the events we observed, i.e., the model's likelihood. Cox proportional hazard log-likelihood is shown in equation(4)<sup>2</sup>.

The log-likelihood can be decomposed in two parts. The first integral corresponds to the probability given by the model to see an event occur, when this event actually occurred, while the second integral corresponds to the probability not to have seen this event before

$$\ell_n(\beta) = \frac{1}{n} \sum_{i=1}^n \left\{ \int_0^{\tau} \sum_{j=0}^p X_{i,j}(t) \beta_j(t) dN_i(t) - \int_0^{\tau} Y_i(t) \exp \left( \sum_{j=0}^p X_{i,j}(t) \beta_j(t) \right) dt \right\} \quad (4)$$

#### 3.3 Time varying variables and coefficients

In video games, variables are changing over time: avatar's characteristics change as the player equips new weapons, difficulty goes slowly up and sometimes peaks up or down. In equation (4) we thus use  $X(t)$ . Our model is then more accurate as we do not summarize  $X$  over time by using a mean for instance, but rather take into account a more timely accurate value of  $X$ .

Variables influence also vary with time. For instance, self efficacy theory states that high levels of difficulty can be more harmful at the beginning of the game. We can model this time varying link between difficulty and motivation with time varying coefficients.

In order to allow coefficient to vary with time without increasing too much the computational complexity, we introduce a piecewise constant estimator. The time interval is discretized in multiple sets, with each set having its own specific coefficient.

Let  $(I_l)_{l \in \{0, L\}}$  be a partition of  $[0, \tau]$ , then

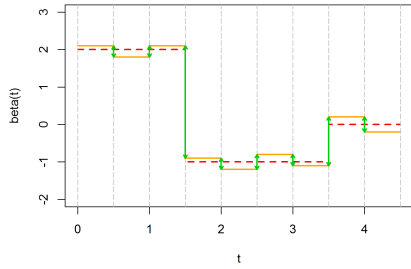
$$\beta_j(t) = \sum_{l=1}^L \beta_{j,l} \mathbb{1}_{(I_l)}(t).$$

To avoid the curse of dimensionality, we introduce a group total variation penalty (5), forcing the model to use as few variations in the parameters as possible. Figure 2 illustrates the effect of total variation penalty. The model without penalty (orange lines) has 9 parameters. The penalty value associated with these coefficients is shown in green. It corresponds to the additional cost related to the combination of successive coefficients. Finally, using a model with total variation (in red) we get a model with only 3 parameters.

This penalty can be combined with a Lasso penalty [19] in order to reduce the number of game variables, excluding those that are less predictive of player retention.

<sup>1</sup>Playtime unit is hours for every figures

<sup>2</sup> $N_i(t)$  is the counting process associated to an event (player departure) and  $Y_i(t)$  equal 1 before the player stop playing and 0 after



**Figure 2: Illustration of Total Variation penalty. Coefficient without penalty (orange) and penalize (red). Penalty value (green).**

$$\|\beta\|_{TV} = \sum_{j=1}^p \left( |\beta_{j,1}| + \sum_{l=2}^L |\beta_{j,l} - \beta_{j,l-1}| \right) \quad (5)$$

As a first step we estimate the parameters that maximize the log-likelihood while minimizing the penalty term. The choice of extra parameter  $\lambda$  is set by cross-validation.

$$\hat{\beta}^{TV} \in \operatorname{argmin}_{\beta \in \mathbb{R}^L} \{-\ell_n(\beta) + \lambda \|\beta\|_{TV}\}, \lambda > 0 \quad (6)$$

Finally, once we get the coefficients estimated with the penalty, we compute the coefficients support. It consists in gathering together close intervals that have similar values. Then we estimate the final coefficients with this new discretization, without penalty.

#### 4 DATA AND GAMES

In this study, we use the previous model on Ubisoft internal data to estimate the link between difficulty and motivation on two games for which we have access to the players' longitudinal data.

The first game is *Rayman Legends*, a platform game developed by Ubisoft and released in 2013. In this game, the players has to jump between platforms to reach the end of a level, by collecting as much as possible of different items and avoiding to be killed by different traps.

In *Rayman Legends*, we consider that the player fails every times he quits a level without finishing it. When the players dies in the level, he spawns to a previous checkpoint, but the death or spawn events are not tracked individually. We thus compute the difficulty as the probability that the player quits the current level.

*Tom Clancy's The Division* is an open world third-person shooter with Role Playing Games mechanics, also developed by Ubisoft and released in March 2016. In this game, the player has to accomplish several missions for his avatar to get stronger, to unlock new regions of the open world and to learn about the game's unfolding story.

In *Tom Clancy's The Division*, we compute the difficulty by considering a failure each time the player is not able to finish a mission, either because his avatar died or because he did not manage to accomplish a specific mission goal and had to try again. Difficulty

estimation is thus a function of the mission id, the difficulty selected by the player for this mission, and various avatar characteristics<sup>3</sup>.

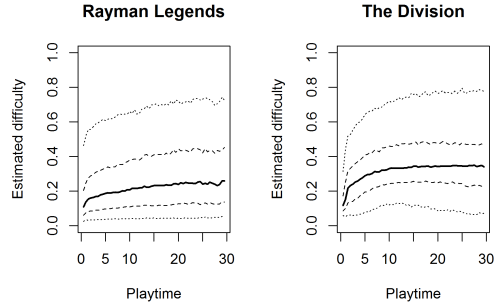
In both games, the sequence of levels is not totally imposed to the player. Several levels are unlocked at specific moments, and the player may always freely choose between different levels, or even to play again a level he already finished.

## 5 RESULTS

### 5.1 Estimated difficulty

For both games we plot estimated difficulty over time using a weighted kernel regression (fig. 3). The area under the ROC curve (AUC) for logistic mixed regression is 80 for *Rayman Legend* and 81 for *The division*. Estimated difficulty in *Rayman Legends* starts around 15 percent and continuously increases to reach nearly 30 percent after 30 hours of playtime. The difficulty smoothly increases as the players level up.

On the other hand, *Tom Clancy's The Division's* difficulty quickly increases during the first game hours, and reaches a plateau of nearly 30 percent. We can notice that difficulty variation among players is higher in *Rayman Legends* than in *Tom Clancy's The Division*, as shown by quantiles lines.



**Figure 3: Estimated difficulty on *Rayman Legends* and *Tom Clancy's The Division*: Median, quantiles 0.25 and 0.75 (dashes) and 0.05 and 0.95 (dots)**

The differences between the difficulty curves can be explained in two ways. First, in *Tom Clancy's The Division*, the player can very easily reduce the difficulty of the game while he may not do so in *Rayman Legends*. In *Tom Clancy's The Division*, the difficulty of a mission depends on the player skills, but also a lot on the characteristics of his avatar. The point is that the player is always able to accomplish easier side missions to buy better weapons, have a more powerful avatar and thus lower the difficulty of the next missions. In *Rayman Legends*, the player may chose between many available levels, but he may not change a level's difficulty. If every open level is hard, his only choice is to beat them.

Moreover, in *Tom Clancy's The Division*, failure is much more punitive than in *Rayman Legends*. When the player attempts a mission, he has to walk from a shelter to the mission site, and missions are often longer than a *Rayman Level*. *Tom Clancy's The Division's*

<sup>3</sup>total health, damages per second for primary weapon, secondary weapon and side arms, stamina, skillpower, electronics and firearms

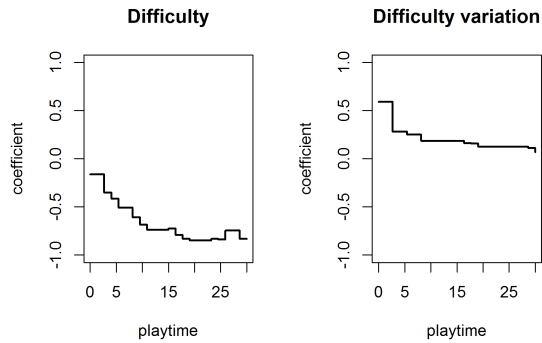


Figure 4: Time-varying coefficient for *Rayman Legend*

slower pace may encourages players to lower their chances of failure and be better prepared for each mission, in order to avoid being frustrated by the time lost. On the contrary, in *Rayman Legends*, the player can restart a level very quickly.

## 5.2 Effect on retention

5.2.1 *Rayman Legends*. First, we studied the effect of difficulty on quit rate, to estimate what level of difficulty fosters the highest level of motivation in players. Results from the time varying Cox proportional hazard model are shown on figure 4 for *Rayman Legends*<sup>4</sup>.

In *Rayman Legends*, we can see that difficulty has a negative link with quit rate. This is equivalent to saying that when they play harder levels, players tend to quit less often. This link between difficulty and retention is growing over time, being close to -0.1 at the beginning and close to -0.8 at the end.

More precisely, the difficulty curve can be interpreted as follows. Say we pick two players A and B that start their 12th hour of playtime. Depending on their skills and the level they choose, difficulty may be at 20% for player A and 40% for player B. The difficulty gap between them is 20%. On the left curve from 4 we can read that the coefficient value associated to difficulty at 12 hours of playtime is around -0.7. Then player A has seen his chances to stop playing multiplied by  $\exp(-0.2 * (-0.7)) = 1.15$  relatively to player B. Reversely player B has 15% more chances to stay in the game than player A.

If causality holds, this results mean that by increasing difficulty, especially in the late hours of the game, we can increase player retention. Moreover, we may give an estimate of the retention gain induced by this change. However, as we are not in a controlled experiment where players are forced to play at a given difficulty level, therefore we can not conclude to causality. Furthermore many unobserved information can affect both difficulty and retention. This experiment shows only interesting correlations that have to be further investigated in controlled experiments to derive a causality link.

Then, we study the effect of difficulty variation on quit rate. Indeed, players may tend to prefer a higher level of difficulty but

<sup>4</sup>Note that we did not plot confidence interval on figures 4 and 5 because they are small and negatively affect readability

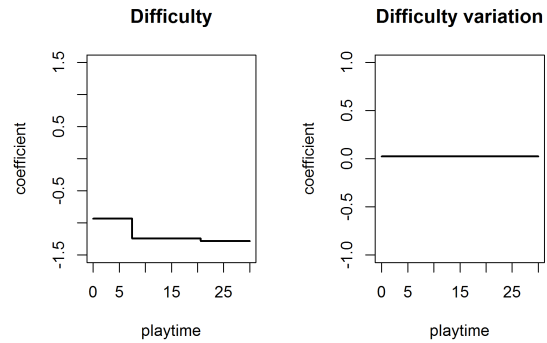


Figure 5: Time-varying coefficient for *The Division*

on the other hand, they may not like to have the difficulty change very quickly from easy to hard, for instance. The right curve on 4 describes the effect of difficulty variation on quit rate. Here, a difficulty variation has to be understood as the difficulty derivative over time.

The results show that in *Rayman Legends*, a high difficulty variation in the first hours of the game will have a negative impact on player retention. Once again if causality holds and given the previous results, a design recommendation could be to smoothly increase the difficulty in the first hours of the game and increase it much more as playtime goes up.

5.2.2 *Tom Clancy's The Division*. Figure 5 show the same analysis on *Tom Clancy's The Division*. Here also we notice than a higher difficulty is correlated with an higher retention. However the effect is strong even during the first hours of the game. In the first hours, a player with a difficulty of 30% has 21% ( $\exp(-0.2 * (-0.95))$ ) more chance to keep playing than a player with a difficulty of 10%. This effect goes up to 27% after 8 hours of playtime.

In *Tom Clancy's The Division*, difficulty variation has no influence on player retention. Even an abrupt increase of difficulty seems to have no impact on player departure rate.

Both these results are coherent with motivational aspects of Video Games that we described in the first sections. First, the flow theory states that players tend to prefer an optimal level of challenge. However, we do not know the sweet spot of such a difficulty level. We may postulate that difficulty is balanced when the player has 0.5 chances to win. If we look at figure 3, we can see that the difficulty is largely distributed under 0.5 for both games. Thus, it seems logical that player tend to prefer higher level of difficulty in both games. Having between 70% and 80% of chances to reach a goal might be considered as lower than what a *challenging* level should be. As Malone puts it, challenge is motivating because it is a source of uncertainty, but it may not be so if the player wins that often.

Then, results in *Rayman Legends* are coherent with the self efficacy theory and Klimmt's experiment. In the beginning of the game, players tend to prefer lower level of difficulty, and do not like the strongest variation of the game's difficulty. Indeed, we can postulate that players need safe practice conditions at the beginning of the game to create a strong belief that they will be able to

beat the game's challenges. Thus, failure by itself, or unpredictable failure due to a quickly heightened level of difficulty at the early stages of the game may be harmful. But after few hours of playtime, players may realize that the game is fair and believe in their own capacity to beat it, and be ready to appreciate more challenging and variables levels of difficulty.

What may be surprising at first glance is that we really do not get the same results with *Tom Clancy's The Division*. In this game, players still tend to link harder levels of difficulty, but this impact of difficulty on retention seems to vary much less over time. Also, it seems that variation of difficulty over time has no impact on player retention. However, we may explain this result through Weiner's Theory of Attribution [23]. In *Rayman Legends*, the main cause of failure is the player's skills. This cause of failure can be controlled, but it has an internal locus of control, and does not change very quickly over time. In *Tom Clancy's The Division*, when the player fails, it is also because of his lack of skills but also because his avatar was not strong enough. Hopefully, this later cause is external (the avatar), and can quickly be controlled by doing some side missions or even more quickly by buying better weapons. This may explain why failure seems to have a much more negative motivational impact on *Rayman Legends* than it has on *Tom Clancy's The Division*.

One last interesting thing to note in *Tom Clancy's The Division* is that, even though players can manipulate the game difficulty explicitly by changing the mission's level, or by increasing their avatar level, results shows that players would like the game to be harder. This may be explained by the fact that avatar progression is not only to be interpreted in term of difficulty : player may prefer to have stronger avatar per se, and regret it later when the games is too easy. Also, as we said in section 5.1, failure is much more punitive in *Tom Clancy's The Division*, and thus, player may be much more cautious when playing to avoid the frustration failure, at the expense of a more balanced level of difficulty.

## 6 CONCLUSION

In this paper, we study the link between difficulty and motivation in video games, using telemetric data from two games developed by Ubisoft: *Rayman Legends* and *Tom Clancy's The Division*. We define difficulty as a probability of failure, that we estimate using a mixed effect logistic regression. We estimate the player's motivation by considering that when his motivation is too low, he will stop playing. We then use survival analysis with time varying variables and coefficients to estimate the link between difficulty and the player's motivation. This is a model we developed and that is fully explained in the first sections.

Our results tend to confirm theories about player motivation. First difficulty is by itself an explanatory variable of player retention. Then, overall, players tend to prefer higher level of difficulty. The flow theory does not state what the optimal level of difficulty is, but in both games, probability of failure is almost always under what we may call a balanced difficulty at 50% chances of failure. More precisely, in *Rayman Legends* player do not want rapid changes of difficulty as well as a higher difficulty for the first few hours. This is coherent with self efficacy theory : failure, at the beginning of the game, can harm the belief in their capacity to success. This point is central to the fact that we do not observe the same results at the

beginning of *Tom Clancy's The Division*. Indeed, in *Tom Clancy's The Division*, causes of failure can be attributed to the avatar strength, which is can be quickly modified. In *Rayman Legends*, failure is mainly due to lack of skills, and thus is much more harmful for the player's motivation during the first hours of the game.

## REFERENCES

- [1] Thibault Allart, Guillaume Levieux, Michel Pierfitte, Agathe Guilloux, and Stephane Natkin. 2016. Design Influence on Player Retention: A Method Based on Time Varying Survival Analysis. In *Proc. IEEE Computational Intelligence and Games Conf., IEEE*.
- [2] Albert Bandura. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84, 2 (1977), 191.
- [3] Z. Borbora, J. Srivastava, Kuo-Wei Hsu, and D. Williams. 2011. Churn Prediction in MMORPGs Using Player Motivation Theories and an Ensemble Approach. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*.
- [4] Kuan-Ta Chen, Polly Huang, and Chin-Laung Lei. 2009. Effect of network quality on player departure behavior in online games. *Parallel and Distributed Systems, IEEE Transactions on* 20, 5 (2009), 593–606.
- [5] Thomas Constant, Guillaume Levieux, Axel Buendia, and Stephane Natkin. 2017. From Objective to Subjective Difficulty Evaluation in Video Games. In *16th IFIP Conference on Human-Computer Interaction (INTERACT)*. Springer International Publishing, 107–127.
- [6] D. R. Cox. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 34, 2 (1972), pp. 187–220.
- [7] Mihaly Csikszentmihalyi and Mihaly Csikszentmihalyi. 1991. *Flow: The psychology of optimal experience*. Vol. 41. HarperPerennial New York.
- [8] Thomas Debeauvais, Cristina V Lopes, Nick Yee, and Nicolas Ducheneaut. 2014. Retention and progression: Seven months in World of Warcraft.. In *FDG*.
- [9] Fabian Hadji, Rafet Sifa, Anders Drachen, Christian Thurau, Kristian Kersting, and Christian Bauckhage. 2014. Predicting player churn in the wild. In *Computational intelligence and games (CIG), 2014 IEEE conference on*. IEEE, 1–8.
- [10] Jesper Juul. 2013. *The art of failure: An essay on the pain of playing video games*. Mit Press.
- [11] Jaya Kawale, Aditya Pal, and Jaideep Srivastava. 2009. Churn prediction in MMORPGs: A social influence based approach. In *Computational Science and Engineering, 2009. CSE'09. International Conference on*, Vol. 4. IEEE, 423–428.
- [12] Jun H Kim, Daniel V Gunn, Eric Schuh, Bruce Phillips, Randy J Pagulayan, and Dennis Wixon. 2008. Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 443–452.
- [13] Nicole Lazzaro. 2004. Why We Play Games: Four Keys to More Emotion Without Story. In *Game Developers Conference*.
- [14] Guillaume Levieux. 2011. *Mesure de la difficulté des jeux video*. Ph.D. Dissertation. Conservatoire national des arts et metiers-CNAM.
- [15] Thomas W. Malone. 1982. Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the 1982 conference on Human factors in computing systems*. ACM, New York, NY, USA, 63–68.
- [16] Torben Martinussen and Thomas H Scheike. 2007. *Dynamic regression models for survival data*. Springer Science & Business Media.
- [17] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion* 30, 4 (2006), 344–360.
- [18] Penelope Sweetser and Peta Wyeth. 2005. GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)* 3, 3 (2005), 3–3.
- [19] Robert Tibshirani. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), 267–288.
- [20] Markus Viljanen, Antti Airola, Jukka Heikkonen, and Tapio Pahikkala. 2017. Playtime Measurement with Survival Analysis. *arXiv preprint arXiv:1701.02359* (2017).
- [21] Guenter Wallner. 2015. Sequential analysis of player behavior. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. ACM, 349–358.
- [22] Ben George Weber, Michael John, Michael Mateas, and Arnab Jhala. 2011. Modeling Player Retention in Madden NFL 11.. In *IAAI*.
- [23] Bernard Weiner. 2005. Motivation from an attribution perspective and the social psychology of perceived competence. *Handbook of competence and motivation* (2005), 73–84.
- [24] Angela Winnett and Peter Sasieni. 2003. Iterated residuals and time-varying covariate effects in Cox regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65, 2 (2003), 473–488.
- [25] Jun Yan and Jian Huang. 2012. Model Selection for Cox Models with Time-Varying Coefficients. *Biometrics* 68, 2 (2012), 419–428.