



HAL
open science

Dynamic Difficulty Adjustment Impact on Players' Confidence

Thomas T.C.C. Constant, Guillaume Levieux

► **To cite this version:**

Thomas T.C.C. Constant, Guillaume Levieux. Dynamic Difficulty Adjustment Impact on Players' Confidence. The 2019 CHI Conference, May 2019, Glasgow, United Kingdom. pp.1-12, 10.1145/3290605.3300693 . hal-02141897

HAL Id: hal-02141897

<https://hal.science/hal-02141897v1>

Submitted on 28 Jun 2019

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.

Dynamic Difficulty Adjustment Impact on Players' Confidence

Thomas Constant

Conservatoire National des Arts et Metiers
CEDRIC
Paris, France
thomas.constant@cnam.fr

Guillaume Levieux

Conservatoire National des Arts et Metiers
CEDRIC
Paris, France
guillaume.levieux@cnam.fr

ABSTRACT

Difficulty is one of the major motivational pull of video games, and thus many games use Dynamic Difficulty Adjustment (DDA) systems to improve the game experience. This paper describes our research investigating the influence of DDA systems on player's confidence, evaluated using an in-game bet system. Our hypothesis is that DDA systems may lead players to overconfidence, revealed by an overestimation of their success chances when betting. This boost of confidence may be a part of the positive impact of DDA systems on the quality of game experience. We explain our method to evaluate player's confidence and implement it into three games related to logical, motor and sensory difficulties. We describe two experimental conditions where difficulty is either randomly chosen or adapted using a DDA algorithm. Results show how DDA systems can lead players to high level of overconfidence.

CCS CONCEPTS

- **Applied computing** → **Computer games**; *Psychology*;
- **Human-centered computing** → **User studies**.

KEYWORDS

User Experience; Video Games; Difficulty; DDA; Dynamic Difficulty Adjustment; Confidence

ACM Reference Format:

Thomas Constant and Guillaume Levieux. 2019. Dynamic Difficulty Adjustment Impact on Players' Confidence. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, May

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHI 2019, May 4–9, 2019, Glasgow, Scotland Uk

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-5970-2/19/05...\$15.00

<https://doi.org/10.1145/3290605.3300693>

4–9, 2019, Glasgow, Scotland Uk. ACM, New York, NY, USA, 12 pages.

<https://doi.org/10.1145/3290605.3300693>

1 INTRODUCTION

Confidence is a complex construct linked to optimism, trust, self-confidence as well as feeling of self-efficacy [51]. In this paper, we study confidence as a feeling of self-efficacy, described by Bandura [9]. Indeed, we will study the impact of difficulty on the player's "*conviction that one can successfully execute the behavior required to produce the outcomes*" [9]. This form of confidence is a desirable positive feeling [31]. Confident players that believes in their chance of success may try-out new and risky strategies and attempt to achieve harder challenges than those being less confident. This might be considered as a useful behavior in the context of safe, virtual environments, and may improve the learning process [1, 42, 50]. In this paper, we show that, as suggested in the literature, the way difficulty is adapted to the player skills can have a strong influence on the player's confidence, leading them in some cases to be highly overconfident.

Our hypothesis regarding the link between difficulty adaptation and player's confidence comes first from the vast literature about the importance of difficulty and challenge in everyday life, as well as in video games. From a positive psychology point of view, Csikszentmihalyi's flow theory argues that quality of experience is impacted by the perceived difficulty of a challenge, and by the perception of one's skills with regard to this challenge [19]. In video games, challenge is often considered as a critical aspect of player's experience. According to Malone, challenge is, with fantasy and curiosity, one of the most important appeal elements of video games [41].

More specifically, flow theory explains that difficulty has to be adapted to the player's skill to provide a motivational boost or to put the player in a flow state. As a practical application of the flow theory for game design, the Gameflow model puts difficulty as one of the criteria of game enjoyment. Gameflow states that levels of difficulty must be varied, and related to the players' skills in order to maintain the players' interest in the challenge [53]. Klimmt et al. also find that too

easy or too hard levels of difficulty lead to boredom or frustration [35]. Even more, the balance or imbalance between player's skills and game's difficulty may provoke different emotional states, a number of them being desirable [45]. Flow, and thus the adaptation of challenge, is a core aspect of video games, but also of computer human interaction in general. As proposed by Elmqvist et al., fluid interaction can be defined as "promoting flow" and thus a "balanced challenge", as well as "supporting direct manipulation" and "minimizing the gulf of action" [22]. Indeed, one may consider that, except in the context gaming, difficulty is to be lowered as much as possible. However, lowering difficulty might also lower the intrinsic value of any interaction. The user may easily lose focus while performing a follow up of trivial task, while they might be more concentrated a flow state is induced by and adapted level of challenge.

Many Dynamic Difficulty Adjustment (DDA) systems have thus been designed to adapt games difficulty in real time, given the player's skills. DDA systems are commonly used within the game industry, whatever the game genre [2, 49], sport [25], shooters [12], platformers [6], role-playing games [24]. DDA systems can for instance be designed to generate an optimal flow experience during the game session [15]. Researchers have investigated their effect on player's enjoyment and motivation [7, 16]. Difficulty and its adaptation is thus a fundamental aspect of game experience.

The link between DDA and players confidence can already be found in the literature, if confidence is seen as a feeling of competence. Flow theory states that participants' sense of competence is ensured by balancing the perceived difficulty of a challenge with the actual skills of the participants [43]. Sense of competence is also examined in Self-Determination Theory as a motivational factor. While applying SDT to games, Ryan et al. study the feeling of competence as a factor underlying players' intrinsic motivation [48]. As video games can provide challenges adapted to players' actual skills, players may feel to be optimally qualified to overcome them and thus, feel a sense of accomplishment.

Confidence can also be linked with the notion of subjective difficulty. Indeed, as stated before, we see confidence as the player's feeling of self efficacy, and thus, of player's confidence towards their future success, which is tightly related to the player's perception of difficulty. Perception of difficulty is fundamental: as Nakamura puts it, the most important aspect of difficulty, when it comes to its impact on game experience, is the player's perception of this difficulty, not the actual one [45]. The thing is that player's perception of difficulty is absolutely not straightforward and can be strongly biased [17]. Furthermore, player's estimation of the useful effort to overcome a challenge depends on multiple factors. The type of challenge, physical or cognitive [21] implies a different

form of effort. Leveux et al. define three dimensions of difficulty, logical, motor or sensory, each one of them requiring various levels of physical and mental effort [8, 38]. Player's expertise and experience may also impact their perception of needed effort [21]. Moreover, Caillois [13], Malaby [40] and Costikyan [18] explain that the amount of effort needed must remain uncertain in order to keep the game compelling. These contributions suggest that one may, by design, have an impact on the player's perception of difficulty, and thus on its confidence.

In this paper, we study how DDA systems can have an impact on the player's confidence. More specifically, we investigate the hypothesis that dynamically adjusting the difficulty according to the player's performance can lead them to overestimate their success chances. We consider that this overestimation is desirable, as it can be thought as a boost of self-confidence. We study the player's confidence while using two DDA systems: one where we pick difficulty uniformly at random, and one that targets a balanced challenge for which players' probability of success equals 50%. We choose to study these two conditions because they are at both ends of the DDA spectrum. Random difficulty does not take into account the player at all, providing him with an unbalanced, unpredictable difficulty, and has thus been used previous research in cognitive psychology about overconfidence [28, 39, 46]. The other one, the fully adaptive DDA targeting $P_{\text{success}} = 0.5$ can be thought as the "canonical" DDA used in video games, where skills and challenge are evenly balanced, as the flow model would suggest. We think that dynamically adapting difficulty to the player's performance will lead them to higher levels of confidence than when difficulty is randomly picked. For these reason, we choose to investigate two DDA systems' impact on the player's confidence, one aiming to *balance* challenge's difficulty, and another one picking *random* difficulty levels.

To do so, we first review various studies related to the use of DDA systems and their impact on the overall player experience, and precisely on the player motivation and enjoyment. Then we discuss the links between game difficulty variations and player confidence. We introduce our method to measure objective and perceived difficulty. Objective difficulty is estimated as the player's chances of success using a mixed effects logistic regression on actual successes and failures. Perceived difficulty is estimated using the player's estimation of their chances of success for a specific challenge using an in-game bet system. We present the design of our games developed to study the influence of difficulty. Each game is dedicated to a type of difficulty: logical, motor and sensory. These games have been played both with a random or balanced difficulty curve, in order to compare these two experimental conditions. Lastly, we present and discuss our

results, showing in which case a balanced difficulty adaptation may have provoked a strong boost of self confidence.

2 DIFFICULTY AND PLAYER'S MOTIVATION

As introduced, flow theory expresses that perceived challenge, as clear goals and feedback, are two conditions for optimal experience [45]. Difficulty, as a challenge parameter, may impact the player's perception of the game experience [35]. Moreover, adapting the objective challenge difficulty to the player's actual skills may help to achieve an optimal experience [15, 53].

Allart et al. study the impact of the difficulty on players' motivation based on two commercial games, *Rayman Legends* and *The Division*. They evaluate the impact of various levels of difficulty on the player retention [5]. In both games, players seem to prefer higher level of difficulty, the average chances of failure being only around 20%. Increasing the difficulty may have a positive impact on the time that the players spent on the games. However, the impact of the variation of difficulty on the players retention is variable from one game to another. Authors postulate that such results may be explained by the nature of the game itself and the perception of the challenge, as related either to player's skills or avatar's levels (locus of attribution). As flow theory states, they find that players favor optimal levels of difficulty, not too easy nor too hard. Players also prefer to face challenge in which the uncertainty about their chances of success is higher.

Alexander et al. test different difficulty settings in order to find the level of difficulty which is the most enjoyable for all kind of players [4]. They show that casual players choose difficulty settings that are not balanced with their actual skills. Casual players seem to prefer lower levels of difficulty, even if their skills are higher than the ones needed by the challenge. On the other hand, experienced players prefer to have difficulty settings accordingly to their skills.

DDA systems may help to address these preferences, as they adapt the difficulty curve for each type of players [8, 30] and, in some cases, according to their affect [3, 47]. Moreover, to balance the player's motivation through the overall game experience, DDA systems may be used in order to maintain a high level of motivation.

Colwell et al. evaluate this adaptive game mechanism by increasing or reducing the number of PNJ assaulting the players during one level of an arcade-style game [16]. They evaluate the enjoyment through a post game questionnaire, crossed with the players' data collected while playing two versions of the same game: one with DDA, one without. Results show that both weak and strong players may enjoy their play experience better when the game difficulty is adapted to their current performance. Players also found their play

experience as more balanced while using DDA systems, according to the post game questionnaire where players have to rate the difficulty on a 1 to 5 scale.

Ang & Mitchell define two types of DDA systems: one based on the players' choices (pDDA), and one other based on their performance (rDDA). They test the impact of these two systems on the nine dimensions of the flow experience [7]. Results show that both DDA systems have a positive impact on the player experience, as they allow players to "enter a state of Flow [...] much faster and for a longer periods of time". They find that pDDA impacts more positively the sense of control, as this DDA system allows the players to control the difficulty of the game. Their study makes a clear link between DDA systems and an important improvement of the flow experience. By extension, immersion and motivation of the players are also impacted by the use of such systems.

Manipulating difficulty with DDA systems has a positive impact on the player's motivation and, according to flow theory, may influence the player sense of competence [20]. Sense of competence may also be linked to other factors, like confidence. For example, cognitive psychology studies show a direct link between the decision-maker confidence and their feeling of self-efficacy [23]. Cognitive biases, such as the *illusion of control* and *overconfidence effect*, may also impact the sense of competence [36]. In this paper, we examine a possible direct link between the use of DDA systems and the player's confidence, as this link has not been studied while playing video games. We make the hypothesis that this connection may partly explain the impact of DDA systems on the players' motivation and enjoyment.

3 DIFFICULTY AND PLAYER'S CONFIDENCE

As introduced, video games are intrinsically built to keep the players highly motivated and engaged during the resolution of challenge. Video games are also used, in specific contexts, as tools to manipulate the player's confidence.

For example, Garris et al. analyze the essential characteristic of video games for educational use. As video games offer experimental environment free from real-world consequences, they can be used to enhance the learner confidence [26]. For the authors, modifying the difficulty of the game may help to increase the players' confidence: "Moreover, games that employ progressive difficulty levels allow the user to gain familiarity and build skills in complex or novel task environments in a graduated manner." In that case, difficulty seems to have a direct impact on the players' confidence.

Confidence is also a key component of the ARCS model, another framework used to describe the game experience. This framework is based on 4 fundamental properties: Attention, Relevance, Confidence and Satisfaction, and originally designed to improve the learners' motivation while facing instructional activities, like games for learning [34]. In this

framework, confidence is described as an expectancy of success, meaning that "[...] *the players have the persuasion to be successful if they show sufficient engagement. The level of confidence has important consequences for several aspects [...] increased and sustained engagement, and self-efficacy.*" [54]. In that case, the perception of the difficulty of the task may be impacted by the learners' level of confidence.

These observations are supported by cognitive psychology literature regarding decision-making and misplaced confidence: "*overconfidence amounts to an 'error' of judgment or decision-making, because it leads to overestimating one's capabilities and/or underestimating an opponent, the difficulty of a task, or possible risks.*" [31]. Such studies demonstrate the presence of a cognitive bias during the estimation of the difficulty of a task called the *hard/easy effect*. This behavior implies that decision-makers cannot estimate the real difficulty of the task [46]. For low levels, they will underestimate their chances of success, whereas for high levels, they will overestimate them [28, 39]. Confidence, in these studies, is defined as the decision-makers ability to estimate their chances of success facing a specific task.

For this research, following cognitive psychology related to overconfidence, we describe the player's confidence as an estimation of their chances of success regarding a specific challenge. In that way, the player's confidence may modify their perception of the challenge, and indirectly, of the perceived difficulty of it. We investigate the effects of DDA systems on the players' overestimation of their chances of success. In other words, adapting the game difficulty to the player's capabilities may induce a feeling of overconfidence. Following that hypothesis, we can make the assumption that the players' motivation may partly rely on a biased estimation of their skills.

Games are tools in which players can experiment various strategies without being confronted to real-world issues [1, 50]. Failure in video games is not definitive but a part of the learning process [27, 33]. Overconfidence in video games, as a misplaced belief in its own capacity of success, may help players to surpass themselves, even if it can lead them to a probable failure. Overconfident players may be more motivated, and subject to a risky decision-making. For instance, a confidence boost may help players to face harder challenges, challenges that are not calibrated on player's actual skills. For educational games, overconfidence may indeed improve the learning process or, at least, arouse learner's curiosity in the challenge. Game designers can also use the player's misplaced confidence to provoke unexpected outcomes, with regard to the unfolding narrative for example. Consequently, studying player's confidence variation in video games may give some methods and provide tools to enhance the overall game experience.

4 EXPERIMENTATION

The current experimentation is investigating a potential influence of DDA systems on players' perception of their chances of success, and thus, on players' confidence. Following the results of our previous research showing the existence of a gap between the objective difficulty of a challenge and the perceived difficulty of the same challenge [17], we want to study the influence of DDA systems on this gap, called the *difficulty estimation error*. As we suggest that balanced difficulty should boost player's confidence, our hypothesis is that the gap is widened when difficulty is adapted to players skills.

In order to investigate these hypotheses, we designed three games, each one of them focused on one of the three dimensions of difficulty defined by Levieux et al.: sensory, logical and motor [8, 38]. Participants have to play each game for 30 turns. At each turn, they have to evaluate their chances of success. One session will use a difficulty curve that follows the player performance and converges to 50% chances of success (*Balanced condition*). To adapt the difficulty, we subtract 0.1 to the difficulty parameter when the player loses, and add 0.1 when they win. By doing so, we adjust difficulty from one step to another based on the player results from the last round. Another session will use a randomly generated difficulty curve (*Random condition*), by picking uniformly at random the difficulty parameter between 0, the lowest level of difficulty, to 1, the highest. Difficulty parameters of each game are described in section 4. The two conditions share the same experimental protocol and context of use. Perceived and objective difficulties are evaluated during the game session, respectively using a bet system and mixed effect logistic regression, as detailed in the following sections.

Measuring Perceived Difficulty

In order to measure players' perception of difficulty, we choose to ask them to estimate their chances of success. If their estimation is high, it means that they evaluate difficulty as low; and if their estimation is low, it means that they think that difficulty is too high for them. But, as found in the literature, this evaluation can be biased by a *hard/easy effect* (section 3. *Difficulty and Player's Confidence*).

Our bet system is based on a 7 point Likert scale (see top of figure 1a) that allows players to bet from 1 to 7 on their success at the current turn of the game. If the player win, the amount they bet is added to their score; if they lose the amount is subtracted. The evaluation of confidence is thus tied to their score, asking them to stay focused and pay attention to such a repetitive question. Betting is done before the players can actually realize their action. To do so, any elements required to solve the challenge is provided, allowing them to evaluate their chances of success.

Our goal is to gather in-game data related to the player perception of difficulty, while avoiding any memory bias that could occur during post performance questionnaires [7, 16, 37]. Measurement of the perceived difficulty is based on the players' bet, noted $D_{perceived}$. With b being the bet value we use the formula $D_{perceived} = 1 - \frac{b-1}{6}$ to get the estimated chances of failure.

Measuring Objective Difficulty

To measure the objective difficulty of a challenge, we follow Leveux et al. definition, suggesting the estimation of the players' failures probability for that challenge [8, 38]. We use a logit mixed-effects model [11] to take into account the players inter-personal differences. Time and difficulty parameter of each challenge (e.g. cursor speed, number of cells...) are used as fixed effect parameters, and we add random intercepts. This mixed model can be used because each player realizes multiple tries on the same challenge. The random intercepts give us a coefficient for each player that we use as a global evaluation of each player's level.¹

Games Description and Difficulty Parameters

Three video games, one for each dimension of difficulty, were played randomly during two sessions. First session was dedicated to the *Balanced* condition, targeting young volunteers, gamers and non-gamers. Second one, carried out one month later, was dedicated to the *Random* condition, targeting the same population with new participants. The experiment was conducted during school vacation period, in a national museum in Paris, the *Cité des sciences et de l'industrie*. We used the same computers and place to fulfill the two experimentation. Each play session was 40 minutes long approximately, including a pre-experiment questionnaire needed to define the different player profile. This questionnaire was split into three parts:

Gaming habits, based on the time that participants spent playing games (board games, video games, social games, gambling).

Self-efficacy profile, based on General Self-Efficacy Scales [10, 14] adapted to gaming situations. This part of the questionnaire is used to check a possible influence of the participants' gaming ability on their estimation of confidence, and by extension, their estimation of their chances of success [52].

Risk aversion profile, based on Holt & Laury's Ten-Paired Lottery-Choice [29]. This part was used to verify any influence of the risk aversion on the players' estimation of their chances of success through the betting system.

¹We also tried to add a random slope on the difficulty parameter but the accuracy gain was very small.

Participants start playing one of the three games, randomly chosen to avoid any order effect. They have to play 33 each game turns, including 3 turns dedicated to the game tutorial. Before starting one game, participants can read its rules and follow their progression through a dedicated page, the game hub. Their goal is to give their best performance, which can be evaluated through their score. All games follow the same protocol: for each turn, player faces a specific challenge for which they evaluate their chances of success by betting from 1 to 7. Then, they can play and see the result. Each game is based on a simple task addressing a dimension of difficulty:

Logical difficulty is evaluated with a sliding puzzle game, where the players have to restore the numerical order of a grid composed of 9 squares (figure 1a). At the beginning of a turn, the grid order is mixed up by switching multiple times the 5th square with an adjacent one. The players have 20 seconds to observe the grid before it disappears. Then they have to make a bet, before replacing the fifth square to its original position, within the number of possible move. The difficulty parameter is associated to the number of possible moves necessary to restore the grid order, linearly from 1 to 11.

Sensory difficulty is evaluated through a perceptive task. A 300 pixels wide grid, composed of multiples squares, is displayed at the center of the screen (figure 1b). At the end of a 3 second countdown timer, five of the squares will blink and then fade out during a short time. The players have to find them back, by clicking on them. The selected squares are displayed in a blue color, whereas the other squares remains in a gray one to avoid any color perception bias. The location of the flashing squares is chosen to avoid simple pattern. The difficulty parameter, d for this game, is related to the fade out time, t . It can be approximated as follows: $t = d^2 - 0.24d - 1.2^2$. The number of squares composing the grid vary with the difficulty: higher levels of difficulty imply that the grid will gain squares on each side, while its surface stay the same. Squares will thus become smaller. The maximum level of difficulty corresponds to a grid composed of 11 squares on each side, while only 4 for the minimum level of difficulty.

Motor difficulty is evaluated with a basic reflex-based task, for which the players have to stop a moving cursor on a black marker at the center of a horizontal segment (figure 1c). The cursor is moving at a linear speed that is correlated with the difficulty of the task, ranging linearly from 100 to 400 pixels per seconds. The sliding area is 320 pixels wide, the cursor is 15 pixels, and the target 2 pixels.

²This equation is a quadratic regression of the fade-out time. In the game, the color is incrementally modified during the game loop, but plotting this equation is much clearer than reading the color update code.



(a) Logical game



(b) Sensory game



(c) Motor game

Figure 1: Interface for the logical (a), sensory (b) and motor (c) games. Logical game is shown with the whole user interface, including the bet system, while we only show the center frame for the other games. Screenshots were taken for the easiest levels of difficulty.

These games share a similar protocol and always provide players with the elements needed to evaluate the difficulty and make a proper bet on their chances of success. For the logical difficulty, the game displays the number of inversions. For the sensory one, the players select the tiles to solve the problem, but without any feedback, before betting. Previous playtests showed that this game could be frustrating if the players had to stop focusing on the grid for betting without selecting the tiles. For the motor difficulty, players can observe the moving cursor before betting.

5 RESULTS

We present here the results of the two experimentation, especially the ones based on the players' in-game performance, as we did not find any correlation between our profiles and estimated player's confidence. The first experimentation, dedicated to the *Balanced* condition, has a total of 6990 observations for 80 participants. The second one, dedicated to

the *Random* condition, allowed to gather 5742 observations for 58 participants. The 138 participants were randomly selected for each condition, and they were not told the purpose of the two experiments. 76 of them declared themselves as regular players. Median age is 15 years old. For each task we remove outliers from the dataset, such as players who did not use the betting system to perform a self-assessment, players who always placed the same bet, or players with outlying performance. For example, a very low score may reflect some user experience issues. Specifically, we have taken out 5 players which had a too low standard deviation on the bets, as they almost always betting the same value. 5 other players were cleared away because they had a too high accumulated bet variation, showing that they took advantage of the adaptive difficulty system in order to maximize their score by deliberately losing with a low bet then by placing a high bet on the next easier challenge and so on. In total, 11 outliers were removed: 6 for the logical game, 4 from the motor game, 4 from the sensory one. Final amount of players for each game is 131 for the logical task, 133 for the motor task and 129 for the sensory task.

Modeling Objective Difficulty

To evaluate objective difficulty, we used a logit mixed effect regression as explained in section 4. *Measuring Objective Difficulty*. Figure 2 describes the conditional R^2 using both fixed and random effects for each game [44]. The model evaluation is performed through a 10-fold cross-validation, using our model as a binary predictor of the challenge outcome. Figure 2 shows that the difficulty parameter is, for all the games, highly significant and has the strongest effect on failure probability, especially for the sensory game. By changing this parameter, we were indeed manipulating the objective difficulty of the games. The effect of time is always negative and significant, meaning that the objective difficulty seems to decrease overtime if the difficulty parameter stays constant. This may indicate that players are learning as their success rate improves overtime for a specific value of the difficulty parameter. We observe that the time effect is stronger for the logical task (-1.51), which is coherent with the fact that players should learn quicker from a logical problem than for a purely sensory or motor one (respectively, -0.46 and -0.27).

For a better understanding of the games' difficulty, we plot the objective difficulty over the difficulty parameter (figure 3). Players are split into three groups of level using k-means³. Figure 3 describes objective difficulty for each challenge for both *Balanced* and *Random* condition at time $t = 0$. The logical game is the most balanced, as the objective difficulty is

³We have players levels from the random intercept of mixed effect logistic regression (see section 4. *Measuring Objective Difficulty*).

Parameters	Logical	Motor	Sensory
Difficulty parameter	5.51 (($p < 2e - 16$) ^{***})	3.07 (($p < 2e - 16$) ^{***})	8.51 (($p < 2e - 16$) ^{***})
Time	-1.51 (($p = 2e - 6$) ^{***})	-0.46 (($p = 0.00018$) ^{***})	-0.27 (($p = 0.0454$) [*])
R^2	0.56	0.28	0.65
Cross Validation	0.71	0.64	0.74

Figure 2: Modeling objective difficulty for each task: logit mixed effect regression results for difficulty and time over failures.

close to the difficulty parameter value. The motor game may be too hard for the lowest levels of difficulty, the objective difficulty is around 0.25 when the difficulty parameter is 0. For the sensory game, it seems that the maximal level of difficulty is reached too quickly, when difficulty parameter is 0.6.

Overall, the objective difficulty model is the weakest for the motor game, with a low conditional R^2 (0.28) and the lowest prediction accuracy (0.64). R^2 and prediction accuracy are higher for the logical and sensory games (respectively, $R^2 = 0.56$ & $accuracy = 0.71$ and $R^2 = 0.65$ & $accuracy = 0.74$).

Differences Between Balanced and Random Difficulty

To investigate a gap between the perceived difficulty of a game using balanced difficulty and one using random difficulty, data for the *Balanced* and *Random* conditions (respectively represented by triangles and dots) are split into 16 equally sized bins using the objective difficulty as estimated by the mixed effect model. In each bin, for each player, the average perceived difficulty is computed. We thus have only one value by player in the bin, and each observation is thus independent from the others. We then only process bins where both *Balanced* and *Random* conditions contain more than 10 different players.

We run a Wilcoxon rank sum test on each bin, against the null hypothesis that perceived difficulty values for the *Balanced* and *Random* condition are drawn from the same distribution ($p < 0.05$), i.e. we check whether perceived difficulty can safely be considered different in the two gameplay conditions for the same objective difficulty level. When the null hypothesis is rejected, triangles and dot are filled, and just outlined when it is not.

Overall results show that players tend to be more overconfident in the *Balanced* condition. For both logical and motor games, the gap is indeed wider for players using a difficulty curve converging to 0.5. Results are particularly strong for

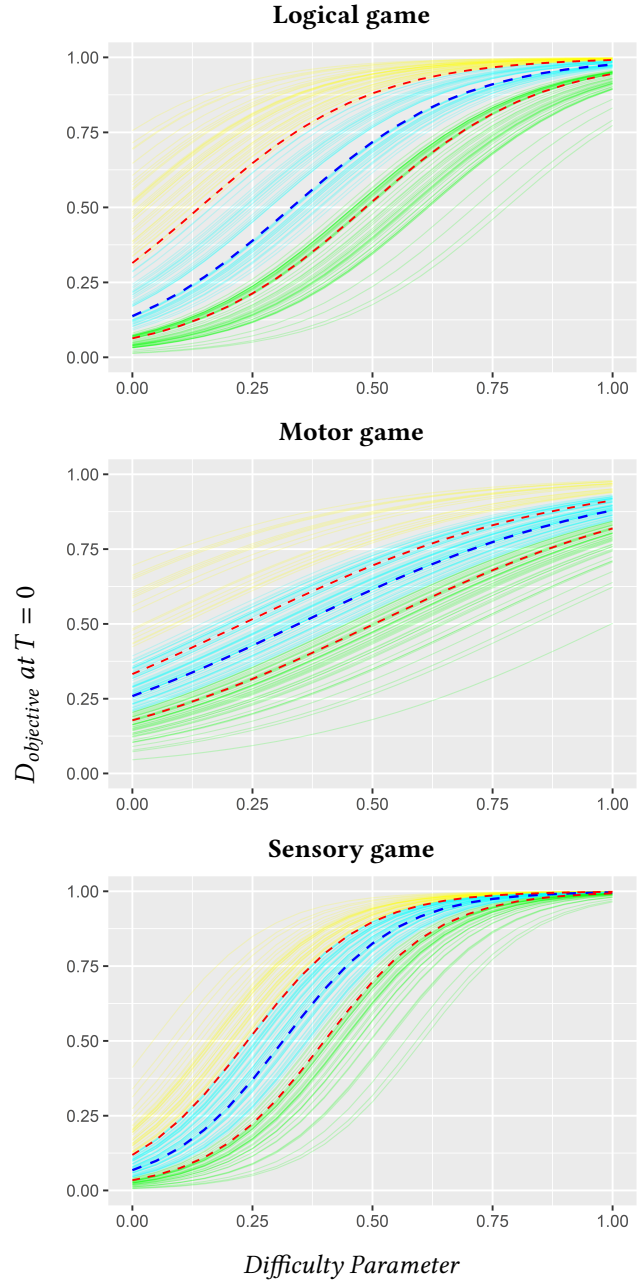


Figure 3: Objective difficulty for each game at $t = 0$, using random and balanced difficulty curves. Blue dashed line is the median player, red dashed lines show first and last quartiles. The less efficient players are in yellow, medium players in cyan and best players in green.

the motor game. For instance, the most significant gap is observed for an objective difficulty at 78% chances of failure: the perceived difficulty for the *Random* condition is at 72% while only at 37% for the *Balanced* condition. Similarly, for

the logical game, when objective difficulty is at 84% chances of failure, the perceived difficulty is at 63% for the *Random* condition, and at 38% for the *Balanced* condition. On the opposite, we do not find any significant results for the sensory game, where it seems that the players have the same slight overconfidence for the two experimental conditions.

Differences Between Objective and Perceived Difficulty

To investigate a gap between the perceived difficulty of a game and its objective one, we run a Wilcoxon signed rank test on each bin, against the null hypothesis that the median of the perceived difficulty values equals the objective difficulty ($p < 0.05$), i.e. we check if perceived difficulty can safely be considered different from objective difficulty. This test evaluates the pseudo-median of the distribution, so we use this value to plot each bin's median on Figure 4. We carry out the test for both the *Balanced* and *Random* conditions, and using triangles for *Balanced* and dots for *Random* condition. When the null hypothesis was rejected and we could safely say that perceived difficulty is different from objective difficulty, we add red diamonds outlines (for *Balanced* condition) and red squares outlines (for *Random* condition) around the corresponding triangles or dots.

Results show that the logical game presents a *hard/easy effect* for both condition. Players seem to be overconfident for the highest difficulty levels, and a little bit underconfident for the easiest difficulty levels. As noted earlier, players of the *Balanced* condition show stronger results. For this experimentation, players of the *Random* condition seem to have a correct estimation of their chances of success for medium difficulty levels included between $D_{\text{objective}} > 0.3$ and $D_{\text{objective}} < 0.7$. Results for the motor game show the strongest hard effect. However players of the *Balanced* condition seem to be overconfident when $D_{\text{objective}} > 0.3$, while players of the *Random* condition seem to be well-calibrated for the same objective difficulty level. For the lowest difficulty levels, only the players of the *Random* condition seem to be slightly overconfident. Players of the sensory game seem to be overconfident for all difficulty levels.

6 DISCUSSION

Balanced & Random Difficulty and the *Hard/Easy Effect*

Each task shows a specific profile with regard to *hard / easy effect*. The logical task seems closest to cognitive psychology studies related to the influence of difficulty on decision-maker's confidence, as we find both an overestimation of the highest difficulty levels (*hard effect*) and an underestimation of the lowest difficulty levels (*easy effect*). However, these effects are only found at the same time in the logical game

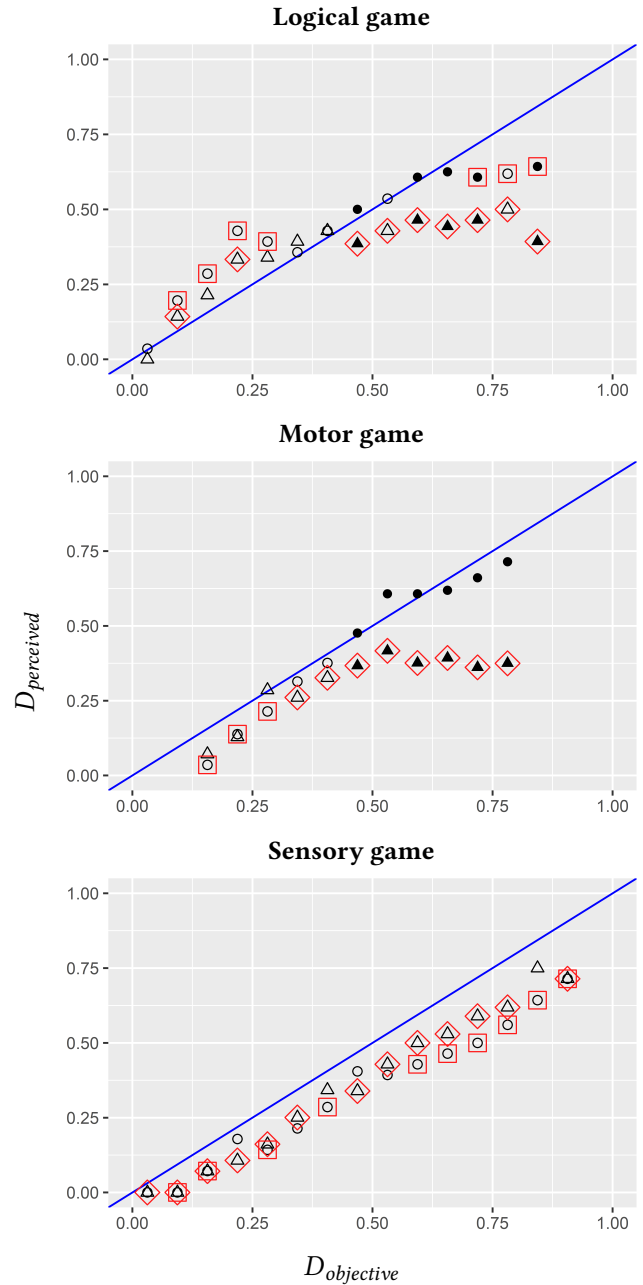


Figure 4: Perceived and objective difficulty for *Balanced* and *Random* conditions, for each game and all players. Triangle represents the players estimation of the difficulty for the *Balanced* condition, and dots for the *Random* one. When triangles or dot are filled, perceived difficulty can be considered different between the two conditions. Red squares and diamonds, respectively for the *Balanced* and *Random* condition, are showed when perceived difficulty is not equal to objective difficulty (blue line).

(figure 4). Following results of previous research [17], we do not find such *easy effect* for the other games. One explanation is that the logical game may be closest to the task used in cognitive psychology studies, implying pure logical challenge like mathematical problems or general knowledge questionnaires.

We observe for the motor game within the *Balanced* condition an important *hard effect*, contrasted by a good estimation of the difficulty for the *Random* condition. Motor game is a simple and repetitive task, meaning that the players' decision-making may be rushed. With random difficulty levels, players may take more time to appreciate the difficulty, and, as such, better evaluate their chances of success.

For the whole experiment, *hard effect* may rely on the player confidence on the game design itself: games are rarely impossible to finish, implying that players think that every difficulty level can be overcome. It may lead to an overconfidence effect. Also, player overconfidence may be stronger in our games than in previous cognitive psychology studies due to player progression. Indeed, our games allow players to experiment and learn from their failures, thereby improving their performance. This feeling of progression and mastery may help players to become more confident in their chances of success. In cognitive psychology studies, where tasks are independent from one to another, this might not be the case.

Differently, results for the sensory game show that players seem to be slightly overconfident for all difficulty levels. One hypothesis is that players perform well and have a better perception of their chances of success because of the nature of the sensory difficulty. Another is related to the sensory game design itself. In this game, the player can select the squares before betting, thus experiencing some gameplay before interpreting their chances of success. While they are not yet aware of their actual performance, they do go one step further toward the completion of the challenge than for the two other games. By assessing their chances of success after having experienced the exercise they may have a more accurate feeling about the quality of their answer. Player's confidence may be reevaluated during this lap of time, canceling any overconfidence effect related to the use of balanced difficulty. For the two other games players perform no interaction and must guess the tasks' next steps. This design choice for the sensory task was made because we did not want to focus on memorization, but on the sensory aspect of detecting blinking squares.

Influence of DDA Systems on the Player's Confidence

We also observe that there is an important difference between the players' estimation of their chances of success depending on how we adapt the difficulty. Players seem to have a better estimation of the difficulty while playing with a random difficulty curve, whereas the difficulty estimation

error is more important when using a balanced difficulty converging to 0.5. In other words, players seem to be much more overconfident with a difficulty adapted to their skills.

As stated in section 2. *Difficulty and Player's Motivation*, DDA systems are used to improve the player's motivation and enjoyment, and our results seem to confirm their impact on the player's confidence. However, one hypothesis could be that when facing unpredictable difficulty levels, players may be more careful with their estimation of their success chances. They know they can be surprised by the difficulty of the next turn and thus not rush their moves and evaluation. This could impact their motivation but others experiments must be conducted in order to verify this hypothesis.

The influence of balanced difficulty on the perceived difficulty seems also to depend on the nature of the task. For the motor game, the gap is wider than for the logical one. Moreover, for the motor game, players seem to have a good evaluation of the difficulty, once again, for the highest difficulty levels in the *Random* condition. Their judgment looks more calibrated than in the other tasks. As we said in section 6. *Balanced & Random Difficulty and the Hard/Easy Effect*, results for the motor game may be explained by the nature of the task, as players may rush their decision-making process. For a pure logical challenge, implying complex judgment and a quite longer decision timing, uncertainty factors are more diverse. Balancing difficulty may thus have a more limited impact on player's overconfidence, as they may tend to be more cautious in their decision process.

We also note that the sensory game presents specific results, in which players seem to be following the same calibration curve for the two experimental conditions. As we said in the previous section 6. *Balanced & Random Difficulty and the Hard/Easy Effect*, the design of the task may help players to have a better estimation of their chances of success, independently from any influence of the DDA system.

7 LIMITATIONS OF THE EXPERIMENT

The Bet System and the Difficulty Evaluation

Our estimation of the challenge perception is based on a bet system, used to measure the difficulty estimation error made by the players. As introduced in section 3. *Difficulty and Player's Confidence*, such a bet system stems from cognitive psychology studies about the overconfidence effect. However, in our games, there exist an optimal strategy that players can use in order to maximize their score: bet 7 when $D_{\text{objective}} > 0.5$, and 1 when $D_{\text{objective}} < 0.5$. Data from players who used this strategy were ditched from our study. We also tried to integrate regular confidence scales in our experimentation, for which players have to give a proper estimation of their confidence. Results show that the players neglect to fulfill this evaluation while it is not strictly related to their score.

Also, the bet system does not allow us to clearly distinguish between effort-based and skill-based perceived difficulties. New experiments can improve the separation between them.

Player Performance and Motivational Influences

Our measurement of players performance is based on their score, without regarding their motivation to exert an effort for one specific challenge. Balanced difficulty seems to boost the players confidence, but we can only make the assumption that such increased confidence may result in a rise of their motivation. If a player is not motivated enough, they may make a correct assessment of difficulty, but perform poorly because they do not want to make the effort. Video game players experience various states of emotion [13, 32], including boredom and anxiety. As such, these emotions should be taken account of for future experiments.

8 CONCLUSION AND PERSPECTIVES

This study investigates the impact of DDA systems on player's confidence. Our work stems from the vast literature describing the link between difficulty adjustment and player experience, from positive psychology studies about flow state, to cognitive psychology studies about overconfidence and *hard/easy effect* and contributions about the impact of DDA systems on player's motivation and enjoyment. Our hypothesis is that DDA systems can place players in a state of overconfidence, that we consider as desirable. To study this hypothesis, we propose a method to evaluate objective and perceived difficulty of a challenge, and implement it into three video games related to various dimensions of difficulty. We report the results of two experiments, where we adapt these games using either a random difficulty curve, or a balanced difficulty curve adapted to players successes and failures.

First, we demonstrate the effectiveness of our objective difficulty estimation, using a mixed effect model to take into account players differences. Results show a predictive accuracy ranging from 64% for the motor game, to 71% for the logical game and 74% for the sensory one. We were also able to see a learning effect, as a negative effect of time on objective difficulty for a given difficulty parameter value. This learning effect is relative to the nature of the tasks, with a higher learning effect for the logical task.

Then, we compare objective difficulty to the perceived difficulty, estimated through a bet system. These results confirm the existence of an unrealistic evaluation of player's actual chances of success when difficulty is balanced with player's skills. In this case, motor game shows a clear *hard effect*, while logical game presents a *hard/easy effect*.

Moreover, when difficulty is randomly chosen, players seem to have a better estimation of their chances of success for the motor and logical game, confirming our hypothesis

for two of the three games. The effect is more important for the motor game, as players of the *Random* condition correctly estimate their odds of win. We also notice the same effect for medium difficulty levels of the logical game. However, for the sensory game only, we observe that players are constantly overconfident, and have a similar perception of the difficulty for both experimental conditions.

We then discussed how the nature of the task may explain these difference. A motor game implies a faster decision-making that may have helped to provoke a high overconfidence state in the balanced condition. The results regarding the sensory task may be explained by the design of this task: the difficulty evaluation is performed after players started the task, thereby potentially gaining additional insight into their performance.

This experiment thus suggests that in the case of motor and logical gameplays, one can influence the player's prediction of self-efficacy by modifying the way difficulty is adapted in real time. Of course, in this experiment, the two conditions used basic difficulty adaptation algorithm : either random or targeting a constant 0.5 success probability. Further experiments needs to be conducted in order to measure the impact of a more various sample of difficulty curves. We want to explore, for instance, the impact of DDA systems targeting higher or lower difficulty levels, or more varied curves with difficulty peaks.

We also plan to investigate DDA systems impact in more complex gameplays. Very often, such gameplays are not based on a turn by turn rhythm. The bet system has to be redesigned in that perspective, as we need to evaluate player's confidence during the game experience without disturbing the player. We want to find a way to keep studying time to time player's confidence, and not rely on post-experiment questionnaires. As the sensory gameplay's result may suggest and as explained in section 4. *Measuring Perceived Difficulty*, timing of confidence evaluation seems crucial.

Finally, we want to expand our approach with an investigation of the impact of these findings on player's behavior. Players showing greater confidence may be disposed to act accordingly. They may attempt harder challenges, try more risky and advanced strategies, or even show an increase of creativity while facing new and unknown challenges. Studying player's in-game actions might help to have a better understanding of the overconfidence effect on the game experience.

ACKNOWLEDGMENTS

This research is part of the *Programme d'investissement d'avenir E-FRAN* project *DysApp*, conducted with *Caisse des Dépôts* and supported by the French Government. The authors would like to thank Laurence Battais & Hélène Malcuit from *Cité des sciences et de l'industrie* for their help.

REFERENCES

- [1] Clark C. Abt. 1970. *Serious Games* (2 ed.). Viking Press, New York, NY, USA. 196 pages.
- [2] Ernest Adams. 2007. *Fundamentals of Game Design* (2 ed.). New Riders, Berkeley, CA, USA. 576 pages.
- [3] Daniel Afergan, Evan M Peck, Erin T Solovey, Andrew Jenkins, Samuel W Hincks, Eli T Brown, Remco Chang, and Robert J K Jacob. 2014. Dynamic Difficulty Using Brain Metrics of Workload. In *CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Matt Jones and Philippe Palanque (Eds.). ACM New York, NY, USA, Toronto, Ontario, Canada, 3797–3806.
- [4] Justin T. Alexander, John Sear, and Andreas Oikonomou. 2013. An investigation of the effects of game difficulty on player enjoyment. *Entertainment Computing* 4, 1 (2013), 53–62.
- [5] Thibault Allart, Guillaume Levieux, Michel Pierfitte, Agathe Guilloux, and Stéphane Natkin. 2017. Difficulty Influence on Motivation over Time in Video Games using Survival Analysis. In *Proceedings of Foundation of Digital Games*. Cap Cod, MA, USA.
- [6] Andy Gavin. 2011. Making Crash Bandicoot: Part 6.
- [7] Dennis Ang and Alex Mitchell. 2017. Comparing Effects of Dynamic Difficulty Adjustment Systems on Video Game Experience. In *CHI PLAY '17 Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM New York, NY, USA, Amsterdam, Pays-Bas, 317–327.
- [8] Maria-Virginia Aponte, Guillaume Levieux, and Stéphane Natkin. 2011. Difficulty in Videogames: An Experimental Validation of a Formal Definition. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, ACE 2011*, Teresa Romão (Ed.). ACM New York, NY, USA, Lisbon, Portugal, 1–18.
- [9] Albert Bandura. 1977. Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review* 84, 2 (1977), 191–215.
- [10] Albert Bandura. 2006. Guide for constructing self-efficacy scales. In *Self-efficacy beliefs of adolescents* (1 ed.), Tim Urdan and Frank Pajares (Eds.). Information Age Publishing, Charlotte, USA, Chapter 14, 307–337.
- [11] Douglas Bates, Martin Mächler, Benjamin M. Bolker, and Steven C. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1 (2015), 1–48.
- [12] Dan Birlaw. 2005. *Resident Evil 4 : official strategy guide*. BradyGames. 223 pages.
- [13] Roger Caillois. 1958. *Les jeux et les hommes : le masque et le vertige* (2 ed.). Gallimard, Paris, France. 374 pages.
- [14] Gilad Chen, Stanley M. Gully, and Dov Eden. 2001. Validation of a New General Self-Efficacy Scale. *Organizational Research Methods* 4, 1 (2001), 62–83.
- [15] Jenova Chen. 2007. Flow in Games (and Everything else). *Commun. ACM* 50, 4 (2007), 31–34.
- [16] Anthony M. Colwell and Frank G Glavin. 2018. Colwell's Castle Defence : A Custom Game Using Dynamic Difficulty Adjustment to Increase Player Enjoyment. In *Proceedings of Artificial Intelligence and Cognitive Science*, John McAuley and Susan McKeever (Eds.). Dublin, Ireland, 275–282.
- [17] Thomas Constant, Guillaume Levieux, Axel Buendia, and Stéphane Natkin. 2017. From Objective to Subjective Difficulty Evaluation in Video Games. In *INTERACT 2017: Human-Computer Interaction - INTERACT 2017*, Regina Bernhaupt, Girish Dalvi, Anirudha Joshi, Devanuj K. Balkrishan, Jacki O'Neill, and Marco Winckler (Eds.). Springer, Mumbai, India, 107–127.
- [18] Greg Costikyan. 2013. *Uncertainty in Games* (1 ed.). MIT Press, Cambridge, USA. 141 pages.
- [19] Mihaly Csikszentmihalyi. 1990. *Flow: The Psychology of Optimal Experience* (1 ed.). Harper & Row, New York, NY, USA. 303 pages.
- [20] Mihaly Csikszentmihalyi. 2014. *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (1 ed.). Springer. 298 pages.
- [21] Alena Denisova, Christian Guckelsberger, and David Zendle. 2017. Challenge in Digital Games: Towards Developing a Measurement Tool. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2511–2519.
- [22] Niklas Elmqvist, Andrew Vande Moere, Hans-Christian Jetter, Daniel Cernea, Harald Reiterer, and TJ Jankun-Kelly. 2011. Fluid interaction for information visualization. *Information Visualization* 10, 4 (oct 2011), 327–340.
- [23] Anders K. Ericsson, Neil Charness, Paul J. Feltovitch, and Robert R. Hoffman. 2006. *The Cambridge Handbook of Expertise and Expert Performance* (cambridge ed.). Cambridge, UK. 918 pages.
- [24] Erik Brudvig. 2010. *Fallout: New Vegas First Look*.
- [25] Phil Frazier. 2008. *Madden NFL 09 Preseason Report*.
- [26] R. Garris, R. Ahlers, and J. E. Driskell. 2002. Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming* 33, 4 (dec 2002), 441–467.
- [27] James Paul Gee. 2007. *Good video games + good learning* (1 ed.). Peter Lang International Academic Publishers, New York, USA. 194 pages.
- [28] Dale Griffin and Amos Tversky. 1992. The weighing of evidence and the determinants of confidence. *Cognitive psychology* 41:1435 (1992), 411–435.
- [29] Charles A. Holt and Susan K. Laury. 2002. Risk aversion and incentive effects. *The American Economic Review* 92, 5 (2002), 1644–1655.
- [30] Robin Hunnicke. 2005. The Case for Dynamic Difficulty Adjustment in Games. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE '05*, Newton Lee (Ed.). ACM New York, NY, USA, Valencia, Spain, 429–433.
- [31] Dominic D. P. Johnson and James H. Fowler. 2011. The evolution of overconfidence. *Nature* 477, 7364 (sep 2011), 317–20.
- [32] Jesper Juul. 2009. *A Casual Revolution: Reinventing Video Games and Their Players* (1 ed.). MIT Press, Cambridge, USA.
- [33] Jesper Juul. 2013. *The Art of Failure* (1 ed.). The MIT Press, Cambridge, USA. 176 pages.
- [34] John M. Keller. 1987. Development of the ARCS model of instructional design. *Journal of Instructional Development* 10, 3 (1987), 2–10.
- [35] Christoph Klimmt, Christopher Blake, Dorothée Hefner, Peter Vorderer, and Christian Roth. 2009. Player Performance, Satisfaction, and Video Game. In *Entertainment Computing ICEC 2009. Lecture Notes in Computer Science*, Stéphane Natkin and Jérôme Dupire (Eds.), Vol. 5709 LNCS. Springer Berlin Heidelberg, Paris, France, 1–12.
- [36] Ellen J. Langer. 1975. The illusion of control. *Journal of personality and social psychology* 32, 2 (1975), 311–328.
- [37] Petri Lankoski and Staffan Björk. 2015. *Game Research Methods: An Overview* (1 ed.). ETC Press. 373 pages.
- [38] Guillaume Levieux. 2011. *Mesure de la difficulté dans les jeux vidéo*. Thèse. Conservatoire National des Arts et Métiers CNAM Paris.
- [39] Sarah Lichtenstein and Baruch Fischhoff. 1977. Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance* 20 (1977), 159–183.
- [40] Thomas M. Malaby. 2007. Beyond Play: A New Approach to Games. *Games and Culture* 2, 2 (apr 2007), 95–113.
- [41] Thomas W. Malone. 1982. Heuristics for designing enjoyable user interfaces: Lessons from computer games. *Proceedings of the 1982 conference on Human factors in computing systems* (1982), 63–68.
- [42] Richard E. Mayer. 2012. *Computer Games for Learning: an evidence-based approach* (1 ed.). MIT Press, Cambridge, USA. 296 pages.

- [43] Giovanni B. Moneta and Mihaly Csikszentmihalyi. 1996. The effect of perceived challenges and skills on the quality of subjective experience. *Journal of personality* 64, 2 (1996), 275–310.
- [44] Shinichi Nakagawa and Holger Schielzeth. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* 4, 2 (2013), 133–142.
- [45] Jeanne Nakamura and Mihaly Csikszentmihalyi. 2014. The Concept of Flow. In *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi*.
- [46] Briony D. Pulford and Andrew M. Colman. 1997. Overconfidence: Feedback and item difficulty effects. *Personality and Individual Differences* 23, 1 (jul 1997), 125–133.
- [47] Pramila Rani, Nilanjan Sarkar, and Changchun Liu. 2006. Maintaining Optimal Challenge in Computer Games through Real-Time Physiological Feedback Mechanical Engineering. *Task-Specific Information Processing in Operational and Virtual Environments: Foundations of Augmented Cognition* 1, September (2006), 184–192.
- [48] 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), 347–363.
- [49] Katie Salen and Eric Zimmerman. 2004. *Rules of play: Game design fundamentals* (1 ed.). MIT Press, Cambridge, USA. 688 pages.
- [50] David Williamson Shaffer. 2006. *How computer games help children learn* (1 ed.). Palgrave Macmillan, New York, USA. 240 pages.
- [51] C. R. Snyder and Shane J. Lopez (Eds.). 2002. *The Oxford Handbook of Positive Psychology* (1 ed.). Oxford University Press. 714 pages.
- [52] Dan N. Stone. 1994. Overconfidence in Initial Self-Efficacy Judgments: Effects on Decision Processes and Performance. *Organizational Behavior and Human Decision Processes* 59, 3 (1994), 452–474.
- [53] Penelope Sweetser and Peta Wyeth. 2005. GameFlow: A Model for Evaluating Player Enjoyment in Games. *Comput. Entertain.* 3, 3 (2005), 24.
- [54] Josef Wiemeyer, Lennart Nacke, Christiane Moser, and Florian Floyd Mueller. 2016. Player Experience. In *Serious Games - Foundations, Concepts and Practice* (1 ed.), Ralf Dörner, Stefan Göbel, Wolfgang Effelsberg, and Josef Wiemeyer (Eds.). Springer International Publishing Switzerland, Chapter 9, 243–271.