

Title	Understanding the Effects of Game in Educational Environment using Game Refinement Measure
Author(s)	Huynh, Phuong Duy
Citation	
Issue Date	2018-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/15198
Rights	
Description	Supervisor:飯田 弘之, 先端科学技術研究科, 修士 (情報科学)

Understanding the Effects of Game in Educational Environment using Game Refinement Measure

By Huynh Phuong Duy

A thesis submitted to
School of Information Science,
Japan Advanced Institute of Science and Technology,
in partial fulfillment of the requirements
for the degree of
Master of Information Science
Graduate Program in Information Science

Written under the direction of
Professor Hiroyuki Iida

March, 2018

Understanding the Effects of Game in Educational Environment using Game Refinement Measure

By Huynh Phuong Duy (1610161)

A thesis submitted to
School of Information Science,
Japan Advanced Institute of Science and Technology,
in partial fulfillment of the requirements
for the degree of
Master of Information Science
Graduate Program in Information Science

Written under the direction of
Professor Hiroyuki Iida

and approved by
Professor Hiroyuki Iida
Associate Professor Kokoro Ikeda
Associate Professor Shinobu Hasegawa
Associate Professor Shogo Okada

February, 2018 (Submitted)

Abstract

In the recent years, there is a growing interest in gamification as well as its applications and implications in education domain because it provides an alternative to engage and motivate students during the process of learning. Whereas gamification is gaining ground in some areas such as business, marketing, and wellness initiatives, applying gamification into education area is an emerging trend. While the majority of studies report overall positive results of applying game-based elements into an education system and evaluating the effects of those elements after applying, the studies, which aim at increasing the effects of those elements in an educational context, are quite restricted. The current problems of increasing effectiveness of game elements in non-game context are it lacks a common measurement and the criteria for assessing that measurement. This is a reason why the authors of gamified systems do not know whether the applied elements make positive effects or not until they do an analysis on users' data and surveys. Moreover, it is really difficult to know exactly what elements and what reasons caused that. Therefore, finding a common measurement and the criteria are needed.

In our studies, we first focused on the entertainment aspect of the gamified system to measure the attractiveness of game elements in this system. We use game refinement theory, which provides a common measurement for quantifying attractiveness of a considered game and be successfully applied into various types of game, in order to measure the attractiveness of game elements in gamification domain. We conduct many analyses based on game refinement measure for understanding more the effect of considered game elements. In the first study, we first focus on analyzing the effectiveness of game element Badge in Duolingo. The result shows that the game refinement values of Badge in popular language courses of Duolingo are lower than sophisticated games. This result is reasonable because Duolingo is a serious environment which means the game elements are not used for making the entertainment environment as fun games. Those values fall into the range between 0.02 and 0.03 which are proposed as suitable range for *GR*-values in gamification applications. Moreover, in this study, a milestone technique was highlighted and its effect also be analyzed. After analyzed, the game refinement trend in the most popular course indicated that the increase of challenge in each milestone aims at adapting the advancement of learners' skill. This trend also express the effect of the course structure that making rest stops for a long learning journey. Moreover, by doing a brief experiment on a course structure, we discovered that Duolingo is enjoyable for new comers who start from the first milestone, however, less enjoyable for advanced users who start at the second or the third milestone.

In the second study, we continue our previous research by doing the comparison between the course structures in different language courses as well as in different applications to see the course structure's effects in the interesting of beginners and the engagement of users for a considered game element. The first result shows that the division and number

of learning material in a structure have an influence on the attractiveness of the considered course. Furthermore, based on the slope of game refinement in the course structure, we have figured out that if the slope is more sloping, users are easier attracted by other game elements. Moreover, the decrease of game refinement value through each milestone can express the decrease of user engagement to the game element. Furthermore, in the comparative study between MindSnacks and Duolingo, the differences of GR -values between these applications contribute to the significance in interpreting learners' enjoyment points. The comparison pattern of GR -value shows that MindSnacks' learners enjoy the use of gamified learning approach rather than Duolingo.

The third study aims at exploring the effects of the game element Winning Streak on the learning process of users in Duolingo. In this research, we first figure out the attractiveness of Winning Streak individually. Then, we discover its contribution in the improvement of entertainment aspect. For the experiment, we collected data of 2000 users to measure a game refinement value of each milestone in the most popular language course. The result of the first consideration has shown that Winning Streak helps users enhance their normal learning activity to serious game activity. That means users is interested their learning activity with Winning Streak. The longer winning streak is more precious for users in Duolingo. After compared the game refinement value trend of Winning Streak and Badge, we recognized that Winning Streak helps Duolingo improve its enjoyment when the attractiveness of Badge decreases. Therefore, the second analysis was conducted in order to clarify this viewpoint. According to the difference of GR -values between two kinds of users, the results show that the streaking-users are more attracted rather than the normal users. Additionally, by comparing the increase ratios of attractiveness and the streaking-users percentages between milestones, the results also expressed that a winning streak is more significant for advanced users rather than beginners

After we understand the effects of game elements and gamification techniques, we apply our experiences and game refinement measure into improving the entertainment aspect of our own game which is cybersecurity awareness training game. This game is designed by using activity theory based-model of serious game and a story game play. The game demo is played and evaluated by 10 participants. The evaluation results indicated that the story game play can help players improve their understanding of cybersecurity problems and resolutions. However, it is rated as not interesting. Therefore, we have used our understanding and game refinement measure to improve the entertainment aspect of this game by designing the learning structure. As the result, a new version of the game has game refinement value falls into the appropriate range of sophisticated games.

Contents

1	Introduction	6
1.1	Background	6
1.2	Problem Statement	8
1.3	Structure of the Thesis	8
2	Game Refinement Theory	10
2.1	Introduction	10
2.2	Basic Idea of Game Refinement Theory	10
2.3	Game Progress Model	12
2.3.1	Time limit sports	12
2.3.2	Score limit sports	12
2.3.3	Board games	13
2.3.4	Appropriate Range of Game Refinement Value in Sophisticated Games	14
2.4	First Application in Gamification Domain	14
3	Gamification Analyses in Language Learning Platforms	16
3.1	Introduction	16
3.2	Language Learning Platform	16
3.2.1	Duolingo	17
3.2.2	MindSnacks	17
3.3	Analysis of Gamification in Duolingo with focus on the Course Structure .	19
3.3.1	Introduction	19
3.3.2	Game Refinement Measure of Game Element Badge	20
3.3.3	The analysis of gamification in Duolingo Language Course	21
3.3.4	Conclusion	25
3.4	The Effect of a Course Structure on the Entertainment in Language Courses	25
3.4.1	Introduction	25
3.4.2	Discovering Effect of Course Structure Based on the Slope of Game Refinement Trend	26
3.4.3	Comparative study: A Case Study in Analyzing Gamification be- tween MindSnacks and Duolingo	29
3.4.4	Conclusion	31
3.5	Analyzing Winning Streak's Effects in Language Course of Duolingo	32
3.5.1	Introduction	33

3.5.2	Winning Streak Attractiveness	33
3.5.3	The contribution of Winning Streak in the Improvement of Entertainment	35
3.5.4	Conclusion	38
3.6	Chapter Summary	39
4	Design and Evaluation of Cybersecurity Awareness Training Game	41
4.1	Introduction	41
4.2	Theoretical Background	42
4.3	Applying ATMSG to Design a Cybersecurity Awareness Game	42
4.4	Evaluation and Discussion	44
4.5	Designing Learning Structure for Improving the Entertainment Aspect	46
4.6	Chapter Summary	47
5	Conclusion	49

List of figures

1.1	Game classification.	7
2.1	Illustration of one level of game tree.	13
3.1	Screenshots of Duolingo.	17
3.2	Screenshots of MindSnacks.	19
3.3	Game refinement value trends when starting from different milestones.	24
3.4	Game refinement value trends of three courses in Duolingo	27
3.5	Game refinement value trends of MindSnacks and Duolingo	31
3.6	Game refinement value trends of Winning Streak and Badge.	36
4.1	Screenshots of CSAG Demo	44

Tables

2.1	Game refinement measures for time limit sports.	12
2.2	Game refinement measures for score limit sports.	13
2.3	Game refinement measures for board games.	14
3.1	Game elements in Duolingo.	18
3.2	Game elements in MindSnacks.	20
3.3	Popular learning languages in Duolingo.	21
3.4	Game refinement measure of Badge in learning language groups.	22
3.5	Number of badges and lessons in EFSS.	23
3.6	<i>GR</i> -values in milestones of EFSS.	23
3.7	<i>GR</i> -values in milestones of Spanish for English Speakers.	26
3.8	<i>GR</i> -values in milestones of French for English Speakers.	26
3.9	<i>GR</i> -values in milestones of Italian for English Speakers.	27
3.10	GR Slope and PoSU in Spanish for English Speakers.	28
3.11	GR Slope and PoSU in French for English Speakers.	28
3.12	GR Slope and PoSU in Italian for English Speakers.	28
3.13	<i>GR</i> -values in milestones of MindSnacks Japanese	30
3.14	<i>GR</i> -values in milestones of Duolingo Japanese	30
3.15	Comparison factors between MindSnacks and Duolingo	32
3.16	<i>GR</i> -values for Winning Streak in milestones of EFSS	35
3.17	Percentage of streaking-users in each milestone.	37
3.18	Game refinement values in detail of each milestone in EFSS.	37
3.19	Real statistic data in EFSS.	38
4.1	Survey questions.	45
4.2	Survey results: average score per question.	45
4.3	Cybersecurity awareness learning content.	47
4.4	<i>GR</i> -values in CSAG	47

Acknowledgment

First of all, I would like to express my greatly appreciate to Professor Hiroyuki Iida for being my great advisor and supporting me many things during my life as a master student in the School of Information Science at Japan Advanced Institute of Science and Technology (JAIST). The door of Professor Iida's office is always open whenever I run into a trouble spot or have a question about my research. He taught me many priceless lessons which is not only limited to research but also how to discover and enjoy the interesting things in life. I have learned many things from him and I sincerely appreciated his favor from the bottom of my heart.

Secondly, I would like to acknowledge Associate Professor Kokolo Ikeda and Associate Professor Shinobu Hasegawa and Associate Professor Shogo Okada for being my committee. I am gratefully indebted to them for their very valuable comments and suggestions on this thesis.

Furthermore, I also would like to say thanks to all other members of my lab for helping me so much during the time I study at the university. Beside these contributions, I would like to thanks my Vietnamese friends, for fulfilling my student life and making me enjoy spending time in Japan.

Finally, I must express my very profound gratitude to my parents and my family for their love and encourage whatever I pursue. Most importantly, I wish to thank my loving and supportive fiancée, Miss Yen Ngoc, who provide unending inspiration and motivation.

Author

HUYNH Phuong Duy

Chapter 1

Introduction

This thesis aims at presenting various analyses of gamification in an educational environment and figuring out the effective use of game elements by considering an entertainment aspect. In the first chapter, we give the background information of our research topic as well as the research question and the structure of this thesis.

1.1 Background

Nowadays, the digital technologies have changed the ways of human working, socializing, communicating, and studying. Game, which has been showcased as one of the digital technologies, is played primarily for the entertainment, especially computer games. It is an important part of leisure lives for the young generation. Along with the development of technology, game has been used to encourage humans to achieve their goal in almost every aspect of life. Game is actually a challenge or a scenario. It presents players with the challenge and forces them to learn new skills and improve abilities. For instance, when a player starts to play tennis, he first tries to hit the ball and miss it. Over the time, after repetition and learning the skills, he becomes a better player. The idea of using games to modify activities is not new. Game has been used to support real-life objectives many times in the past. For example, sports are used for motivating humans to do exercise and to make healthy habits, or simulation games are used for training and enhancing players' skills. In the recent years, this idea has become especially popular.

In education, game has been integrated in order to form an innovative educational paradigm [31]. Game can be used for fostering a learning process effectively and interestingly. Applying game into an educational system takes the advantages of gaming technologies to create an enjoyment, motivation, engagement, and enhancement in learning [12]. The benefit of game and game-based approaches in education domain has been investigated since the 1980s [6]. Many researchers believe that the approaches can better motivate entertainment-driven learners to more thoroughly engage in learning through meaningful activities, which are defined in the educational context, as opposed to those offered using more traditional didactic approaches.

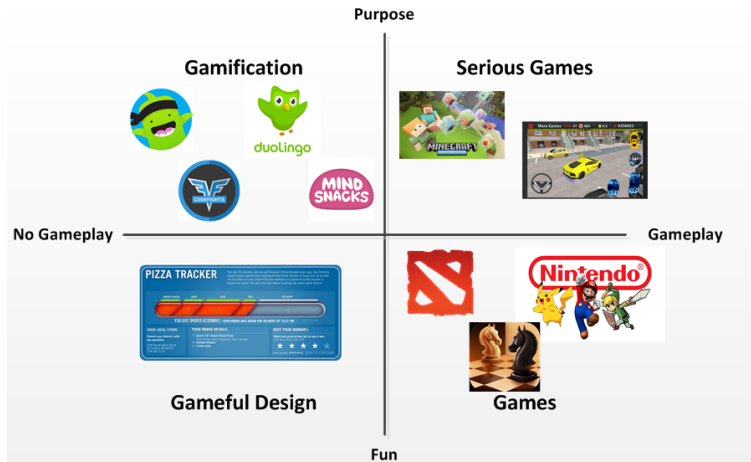


Figure 1.1: Game classification.

According to the classification of game in Figure 1.1, the use of game for a serious purpose is categorized as serious games and gamification. The term gamification refers to the use of game-based elements such as game mechanics, aesthetics, and game thinking in non-game contexts. Gamification as a term originated in the digital media industry. The first documented uses dates back to 2008 [20], but gamification only entered widespread adoption in the second half of 2010 [22], when several industry players and conferences popularized it. Whereas gamification is gaining ground in some areas such as business, marketing, and wellness initiatives, its applications in education area is still an emerging trend [32]. Unlike serious games, gamification applications do not have a game play designed for a specific purpose. It is only the gather of game elements which are used for different purposes without change a practice of existing system. Therefore, some famous companies such as Amazon and Foursquare have used game elements for gamifying their system in order to attract and engage their users.

Most of the recent studies in the domain of gamification, especially in educational environment, aim to figure out the effective way for applying game elements into their learning system. Thus, many gamification design models and frameworks [2] [13] [10] are figured out for pursuing that target. However, almost gamification products are assessed via a survey or usability test [15] [25]. Those assessment methods spend a lot of resources and only give general results. The authors of gamified systems do not know whether the specific applied game elements make positive effects or not until they conduct an assessment. In case the use of game is not effective, it is really difficult to know exactly what game elements and what reasons caused that. Therefore, our works aim at finding a common measurement which reflects an effectiveness or entertainment aspect of game elements. From this measure, we can assess and analyze effects of the particular game element. According to conducted analyses, we can give an effective strategy for using game elements in an existing system.

1.2 Problem Statement

As we have mentioned, our research aim at optimizing the use of game elements in an educational environment. Therefore, the problem statement is expressed below.

Problem Statement How to use a specific game element effectively?

In educational games, the purpose of those games is twofold: (i) to be fun and entertaining, and (ii) to be educational. Therefore, assessments of educational games must be considered on two aspects which are entertainment and educational impact. In our works, we try to analyze and assess some popular game elements which have been used in popular gamified systems. We first aim at figuring out an assessment method for the entertainment aspect which can give a common measurement for effectiveness of game element. The method for our assessment is figured out from game refinement theory idea. The game refinement theory is a unique theory which shows a particular way to quantify the attractiveness and sophistication of the game under consideration. This theory was proved and applied by several studies [16] [17]. Based on the measure, many analyses are conducted in order to find the best reasonable answer for our research question. In this thesis, we use some popular language learning gamified platforms as our testbeds such as Duolingo [9] and Mindsnacks [24]. Users data in those platforms is also collected and used to make a comparison with our analyses.

1.3 Structure of the Thesis

The structure of this research contains five chapters.

Chapter 1 is the introduction. In this chapter we give some background information about gamification and some recent studies as well as the research trend of this domain. We also express our problem statement and a research plan for solving the research problem.

In Chapter 2, we provide background knowledge of game refinement theory which is the main theory used in this research. Their recent works and our idea for applying them into our research domain are also presented in this chapter.

In Chapter 3, we present various analyses of game elements in two language learning platforms Mindsnacks and Duolingo. In this chapter, we introduce the use of game refinement measure for considering entertainment aspect. We also indicate our point of view about using game in educational environment

Chapter 4 presents the design of cybersecurity awareness training game. The design idea of this game and its assessment are indicated this chapter. We use our understanding about game elements and game refinement idea in the previous chapter for improving

an entertainment aspect of our games as well.

Finally, Chapter 5 provides the summary of this research and our future works.

Chapter 2

Game Refinement Theory

This chapter aims at giving a review of game refinement theory [30]. This theory is a unique theory which shows a particular way to quantify the attractiveness and sophistication of the game under consideration. Invented by Iida *et al*, the game refinement theory was proved and applied by several studies [16] [17] [36] .

2.1 Introduction

Classical game theory [11] originated with the idea of the existence of mixed-strategy equilibrium in two-person zero-sum games. It has been widely applied as a powerful tool in many fields such as economics, political science, psychology, and computer science. Different from game theory, which concerns the winning strategy from the player's viewpoint, the game refinement theory aimed at concerning the entertainment impact from the game designer's point of view. This new game theory, which was proposed based on the concept of game outcome uncertainty, focus on the attractiveness and the sophistication of games. The foundation of this direction has been done by Iida *et al* [16].

In the beginning, a logistic model was constructed in the framework of game refinement theory and applied to many board games including chess variants and Mah Jong [17]. Recently, a mathematical model of game refinement has been proposed based on game information progress. After that, it also has been applied in various types of sports games [27] and video games [7]. Further investigation of game refinement theory has been made in various ways. An important way is to figure out a reasonable model of the game progress in the game under consideration.

2.2 Basic Idea of Game Refinement Theory

A mathematical model was constructed based on the concept of game information progress [30]. It bridges a gap between board games and sports games. Firstly, we describe a mathematical model of game progress in order to derive a game refinement measure (*GR*-value). Then, we present its previous applications in fun game domain such as board games, time

limit sports, and score limit sports. The new application of game refinement theory in gamification domain for considering entertainment aspect is also expressed after all.

The game progress is twofold: **(i)** is a game speed or scoring rate and **(ii)** is a game information progress with focus on the game outcome. Game information progress represents the degree of certainty of game's result in time or in steps. Having full information of the game progress, i.e. after its conclusion, game progress $x(t)$ will be given as a linear function of time t with $0 \leq t \leq t_k$ and $0 \leq x(t) \leq x(t_k)$, as shown in Eq. (2.1).

$$x(t) = \frac{x(t_k)}{t_k}t \quad (2.1)$$

However, the game information progress, which is given by Eq. (2.1), is unknown during the in-game period. The presence of uncertainty during the game, often until the final moments of a game, reasonably renders game progress as exponential. Hence, a realistic model of game information progress is given by Eq. (2.2).

$$x(t) = x(t_k)\left(\frac{t}{t_k}\right)^n \quad (2.2)$$

In Eq. (2.2), n stands for a constant parameter which is given based on the perspective of an observer of the considered game. Only a very boring game would progress in a linear function. Therefore, it is reasonable to assume a parameter n , based on the perception of game progress prior to completion. If the information of the game is known completely (i.e. after the end of the game) and the value of n is 1, the game progress curve appears as a straight line. In most games, especially in competitive ones, much of the information is incomplete, the value of n cannot be assumed. Therefore, game progress is a steep curve until its completion, along with $x(t_k)$, t_k , $x(t)$, and t , just prior to game's end. Then acceleration of game information progress could be obtained by deriving Eq. (2.2) twice. Solving it at $t = t_k$, we have Eq. (2.3).

$$x''(t_k) = \frac{x(t_k)}{(t_k)^n}(t_k)^{n-2}n(n-1) = \frac{x(t_k)}{(t_k)^2}n(n-1) \quad (2.3)$$

It is assumed in the current model that game information progress in any type of game is encoded and transported in our brain. We do not yet know the mechanism of information processing in the brain, but it is likely that the acceleration of information progress is subject to the forces and laws of physics. Too little game information acceleration may be easy for human observers and players to compute and becomes boring. Contrary, too much game information acceleration surpasses the entertaining range and will be frustration, and at some points beyond that could become overwhelming and incomprehensible. Therefore, we expect that the larger the value $\frac{x(t_k)}{(t_k)^2}$ is, the more exciting the game becomes.

Thus, we compute its root square, $\frac{\sqrt{x(t_k)}}{t_k}$, and use the result as a game refinement measure for the considered game. We assign it as R -value, which is short shown in Eq. (2.4).

$$GR = \frac{\sqrt{x(t_k)}}{t_k} \quad (2.4)$$

Sports	G	T	GR
Soccer	2.64	22	0.073
Basketball	36.38	82.01	0.073

Table 2.1: Game refinement measures for time limit sports.

2.3 Game Progress Model

This is a significant and challenging task to figure out the reasonable model of game progress and game information progress of a game under consideration. Three subsections below show how the game progress model was figured out in various game domains such as board games, time limit sports, and scored limit sports. The game refinement measure of some popular games as well as its appreciate range in fun games domain is also indicated in those subsections.

2.3.1 Time limit sports

In previous study [30], the mathematical model of game progress is figured out and applied first in time limit sports games such as Soccer and Basketball. In sports games, the scoring rate is calculated by two factors which are (i) goal and (ii) time or steps to achieve the goal. For example, in Soccer, the goal of this game is given by the average number of successful shots whereas the steps to achieve the goal is estimated by the average number of shot attempts. Therefore, the game refinement measure GR is calculated as shown in Eq. (2.5) where G and T stand for the average number of successful shots and the average number of shot attempts respectively. The values G and T correspond to $x(t_k)$ and t_k which are discussed in the section 2.2. The GR -value of some popular games in this domain is shown in Table 2.1

$$GR = \frac{\sqrt{G}}{T} \quad (2.5)$$

2.3.2 Score limit sports

In score limit sports, there is no time limit in those games. The game is regulated by a score limit. Therefore, the game progress model of those games is constructed with focus on two factors W and T which stand for the average scores of a winner and the average total scores of an entire game respectively. Hence, the game refinement measure GR is calculated as shown in Eq. (2.6). The values W and T correspond to $x(t_k)$ and t_k in the previous discussion. The results from previous studies are indicated in Table 2.2.

$$GR = \frac{\sqrt{W}}{T} \quad (2.6)$$

Sports	Version	W	T	GR
Volleyball	Side-out system (15pts)	15	52.52	0.074
	Rally point system (30pts)	30	53	0.103
	Rally point system (25pts)	25	44	0.114
Badminton	Old scoring system	30.07	45.15	0.121
	New scoring system	46.34	79.34	0.086
Table-tennis	Pre-2000	57.87	101.53	0.075
	Post-2000	84.86	96.47	0.077

Table 2.2: Game refinement measures for score limit sports.

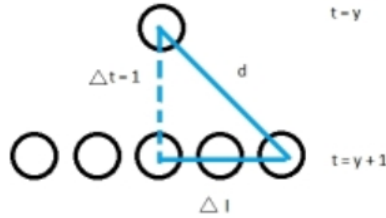


Figure 2.1: Illustration of one level of game tree.

2.3.3 Board games

We consider the gap between board games and sports games by deriving a formula to calculate the game information progress of board games. Let B and D be the average branching factor (number of possible options) and game length (depth of whole game tree), respectively. One round in board game can be illustrated as a decision tree. At each depth of the game tree, one will choose a move and the game will progress. According to Figure 2.1, the distance d , can be determined by using simple Pythagoras theorem. As a result, we have $d = \sqrt{\Delta t^2 + 1}$.

Assuming that the approximate value of horizontal difference between nodes is $\frac{B}{2}$, then we can make a substitution and get $d = \sqrt{(\frac{B}{2})^2 + 1}$. The game progress for one game is the total level of game tree times d . For the meantime, we do not consider Δt^2 because the value ($\Delta t^2 = 1$) is assumed to be much smaller compared to B . The game length will be normalized by the average game length D . Then the game progress $x(t)$ is given by $x(t) = \frac{t}{D}d = \frac{t}{D}\sqrt{(\frac{B}{2})^2 + 1} = \frac{Bt}{2D}$. Therefore, in general, the game refinement measure GR for board games is shown in Eq. 2.7. Iida *et al* [16] calculate the game refinement values for various board games such as chess, Go, and Mah Jong [17]. We show, in Table 2.3, the results.

$$GR = \frac{\sqrt{B}}{D} \quad (2.7)$$

Game	B	D	GR
Western chess	35	80	0.074
Chinese chess	38	95	0.065
Japanese chess	80	115	0.078
Go	250	208	0.076
Mah Jong	10.36	49.36	0.078

Table 2.3: Game refinement measures for board games.

2.3.4 Appropriate Range of Game Refinement Value in Sophisticated Games

As many previous studies confirm, it is obvious that game refinement theory can be effectively applied in many domains of game such as board games [17], sports games [27], video games [7], and even educational board games [19]. This theory can be used as a helpful tool to quantify the attractiveness of a game and it also enables game designers to make a target game more sophisticated by adjusting a game setting or game rules. For example, according to the analysis of volleyball in Table 2.2, game designers will know which version or which scoring system is able to make players are more excited. However, higher GR -value does not mean a game is better. If GR -value is too high, which means players feel too excited any time they play, players will become saturated after several times. On the other hand, if GR -value is too low, which means the game is too boring at the beginning times, players will fall into frustration quickly. Therefore, the game refinement measure has its own appropriate range. Players will be comfortable while playing the games which have the GR -value in that range. As a tentative conclusion, we observed that suitable range of game refinement value is around 0.07 – 0.08, with many previous studies confirmed.

2.4 First Application in Gamification Domain

Educational games are different from fun games. They are the gamified learning platforms, which means those game elements are used to create enjoyment points in learning. As we have stated, one or a bundle of game elements is used for a specific purpose. For this purpose, a reward or an achievement such as certificate, badge, trophy, rank, or title is given to users for encouraging them to complete the author’s tasks. Hence, an attempt to complete learning tasks for obtaining rewards is considered as a game action in educational environment. Therefore, the game progress model has two factors R and T which represent the average number of achievements or rewards and the average number of efforts or tasks respectively. The values R and T correspond to $x(t_k)$ and t_k in the previous discussion and the general formula for calculating game refinement measure GR is expressed in Eq. 2.8.

$$GR = \frac{\sqrt{R}}{T} \quad (2.8)$$

Chapter 3

Gamification Analyses in Language Learning Platforms

3.1 Introduction

The continually increasing number of language-learners over the world has caused the appearance of many tech-based platforms for learning other foreign languages. Many practices of learning languages via websites, various software, and mobile applications have recently been becoming a hot trend due to the emerging of digital technologies. These varieties of learning platforms have been making a lot of changes towards the style of language-teaching and learning as compared to traditional approach. In educational games, gamification has become a prominent word in consolidating intrinsic motivation with such elements like challenge, curiosity, control, and fantasy [26]. Those elements listed are having their own special strategies to engage with learners.

According to the book of UDL Technology [28], MindSnacks and Duolingo have been awarded as “Honorable Mention”. This award is “Best of Class” award for the good technologies. These applications have been using gamification in their learning languages system as compared to non-gamified platform such as Babbel, Busuu. That means gamified learning applications are top choices for learning language users. Therefore, in this research, we do some analyses to discover and explain how a game element can attract learners by using a game refinement measure. Our experiments are conducted on Duolingo and MindSnacks. The analyses are also compared with real user data for proving our hypothesis.

3.2 Language Learning Platform

In this section, we aim at giving the background information of two popular learning language platforms which are MindSnacks and Duolingo. The given information includes basic information, the popularity, and used game elements’ information in an introduced platform.

3.2.1 Duolingo

Duolingo is a gamified free language-learning platform which is developed by professor Luis Von Ahn and his graduate student Severin Hacker. This platform is designed so that users could learn languages while helping Duolingo to translate documents. Duolingo became publicly available in 2012 with more than 300000 users. As of April 2016, it offers 59 different language courses across 23 languages [9]. This platform has quickly become one of the most popular ways for learning foreign languages on the Internet.

Originally, Duolingo aims to provide a free and enjoyable global language learning platform and was intended to be enriched by the translations of its learners. After that, Duolingo began providing a digital language certification program Test Center, which is intended to serve as an alternative to the Test of English as a Foreign Language (TOEFL iBT) and other language proficiency tests. The authors of Duolingo skillfully use some game elements in their platform in order to engage and motivate users. Those game elements include Experience Point (XP) and Level, Badge, Leader Board, Winning Streak, Golden Badge, and Lingot. The detail of those elements are shown in Table 3.1.

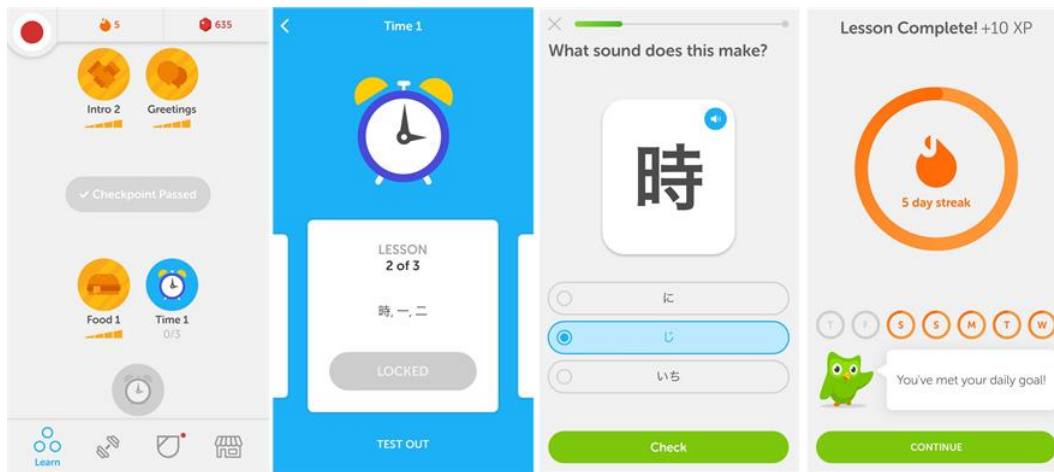


Figure 3.1: Screenshots of Duolingo.

3.2.2 MindSnacks

Unlike Duolingo which is designed as a quiz game, MindSnacks' authors have used several mini games for attracting users in learning languages. This application is available mainly on iOS mobile platform. In 2012, MindSnacks raised 6.5 million dollar from Sequoia with a core mission which is to help gamers turn those nagging Angry Birds and Fruit Ninja addictions into opportunities for learning. Now, this application offers 9 different language courses for English speakers. MindSnacks provides 9 additive and unique games for practicing a learning language in many areas such as vocabulary, grammar, and spelling. Each game is designed with a personalized learning algorithm that helps users maximize

Game element	Obtaining way	Purpose	Description
XP and Level	It is a reward for completing certain activities such as completing a lesson and practicing.	Providing a way to track users' daily activities and compare with other users.	XP determine the language Level, which is displayed on a profile page. Level is a measure of how much work user have put into the language.
Badge	This is a reward for completing all lessons in skill.	Motivating users to complete their lesson.	There are several types of skills used to teach concepts to language learners. Each skill contains a certain number of lessons. Badges are awarded after user complete to learn these skills.
Leader Board	User's rank will increase if their XP pass over this of their friends.	Creating motivation by making a competition among users.	Leaderboard shows users how they are doing compared to their friends, these are some great motivators to help the users coming back and learning.
Winning Streak	Users must meet their XP goal before midnight.	It encourages users to learn a language every day if they want to extend their streak's length.	Streak is a measure of how consistently the users use Duolingo. It starts at zero and increases by one for each day the user meets their XP goal. It resets to zero when the goal is missed.
Golden Badge	Users strengthen their weak skills.	It helps users: memorize, test their knowledge, review learned material, and maintain their skills.	According to Duolingo, learners should first focus on finishing all the lessons in a skill, and later come back to review them at just the right times.
Lingot	Lingots can be earned in several ways: completing a skill in the first time, reaching the next Level, maintaining a streak.	This is an extra reward which helps users increase their motivation in learning activities.	This is a virtual currency in Duolingo which is used to buy some support tools in Duolingo Store.

Table 3.1: Game elements in Duolingo.

memorization, retention and contextual usage at their own individualized pace. Collectively, those games were played while learning 50 different lessons that teach grammars and contextual concepts along with over 1,000 essential vocab words, and phrases. In MindSnacks, the authors also used some game elements for motivation and engagement in learning such as XP and Level system, Mini games, and Hot Streak (see Table 3.2).

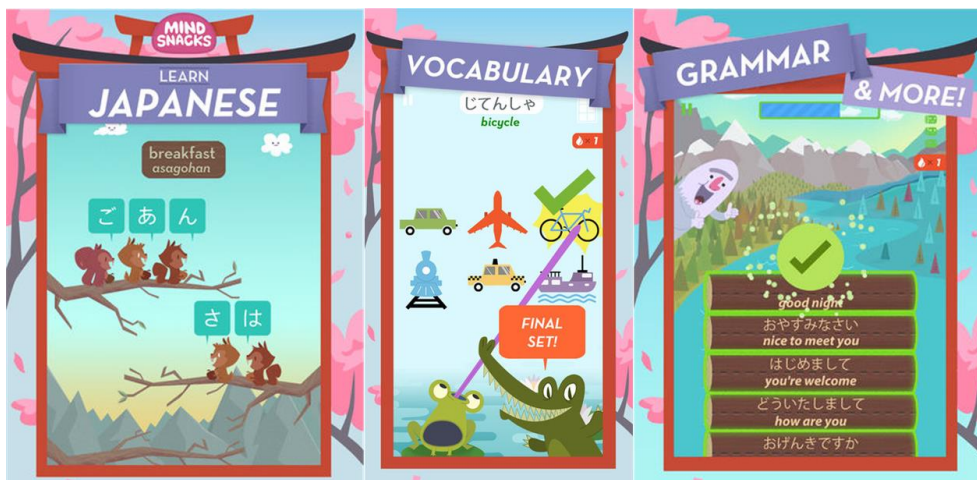


Figure 3.2: Screenshots of MindSnacks.

3.3 Analysis of Gamification in Duolingo with focus on the Course Structure

3.3.1 Introduction

We first start with the analysis in Duolingo. Although Duolingo has used many game elements in their platform, we especially start to analyze Badge in our first study because the important learning activity in Duolingo is to acquire new knowledge by completing a new lesson. For encouraging users to complete a learning lesson as much as possible, Badge is given to lift up the motivation of learners when they study. Moreover, Badge is combined in harmony with a learning content to construct the main structure of a language course. The structure of a language course includes some elements as follows. The core element in a course is learning lessons. A lesson is well-designed, drilling skills of user with several different kinds of challenges. Some lessons are categorized into a small set, which is called a skill, by vocabulary meaning such as verb, adjective, sport, food, etc. Each skill has a strength bar, which will be full only when users have passed all lessons in the skill. At the beginning, only basic skill is available. Other skills are locked until users complete all available skills. The skills in a skill-tree are split into several milestones which present for each stage of user's study process. Although there is no reward given when users reach each milestone, the milestone technique helps users split a boring and challenging learning process into several shorter and easier processes. That helps users

Game element	Obtaining way	Purpose	Description
XP and Level	It is a reward for completing lesson and quest.	Providing a way to track users' daily activities and used for level upping to unlock a new mini game. It makes motivation in learning.	XP determine the user's Level. Level is a measure of how much work user have put into the language. For unlocking a new mini game, user is required to reach a certain number of levels.
Mini games	A new mini game is obtained by reaching a required level.	Making entertainment and motivation in learning language.	A different mini game has different game play. Each mini game is used for practicing each area in learning language. Users can play any available mini game to learn any lessons.
Hot Streak	Users are rewarded for completing a quest.	It encourages users to learn language following a given plan in a learning quest.	Streak is a measure of how consistently they complete a learning quest.

Table 3.2: Game elements in MindSnacks.

feel accomplished and refreshed after learning a ton of new knowledge and they are more motivated and go further in their learning process. Therefore, in this study, we especially focus on the course structure and its game element-Badge for analyzing the gamification of Duolingo.

3.3.2 Game Refinement Measure of Game Element Badge

For analyzing the effect of Badge on the entertainment aspect, we use the game refinement measure GR as the attractiveness caused by Badge. The analyses and interpretations are conducted based on the value and the change of this measure in considered environment. According to the general formula of GR mentioned in 2.4, we derive the formula for calculating game refinement measure by detecting two parameters R and T . In the situation where users aim at obtaining new knowledge, a badge, which is given to user after they complete all lessons in skill, is considered as a reward. Therefore, the average number of obtained badges B and the average number of learned lessons L replace R and T in the formula correspondingly. As a consequence, we have an Eq. 3.1 used to calculate GR -value for game element Badge.

$$GR_{Badge} = \frac{\sqrt{B}}{L} \quad (3.1)$$

3.3.3 The analysis of gamification in Duolingo Language Course

In this research, we conduct two analyses for two purposes. The first analysis aims at analyzing the attractiveness of game element Badge in all popular language courses in Duolingo. While another focuses on analyzing a course structure in the most popular language course “English for Spanish speakers”(EFSS).

Attractiveness of Badge

Duolingo is a learning language application where users cannot ‘lose’ in learning as playing a fun game. Therefore, in a particular language course, enrolled users will obtain the same certain number of badges and learn the same certain number of lessons. That means the average number of obtained badges B and the average number of learned lesson L are exactly the number of badges and number of lessons in considered course. To collect data for B and L , we access to the home page of Duolingo language course. The number of badges is presented on the skill-tree and the number of lessons is counted by assessing to each skill (see the first and the second picture in Fig. 3.1).

Duolingo has 120 million users around the world and currently teaches 19 distinct languages. The most popular courses are available for speakers of a variety of languages. For example, we can learn English from 21 different languages [29]. In order to see an effect of Badge, we collect data in popular language courses which are shown in Table 3.3. For measuring the effect of game element Badge in learning of a particular language, we group all courses which have the same learning language together. For example, the course “English for Spanish speakers” and “English for French speakers” are in the same group.

After data was collected, we calculate the game refinement measure in each language group by using the formula in Eq. 3.1. For instance, we have to find the average number of obtained badges and the average number of learned lessons of 21 courses in English group for calculating its game refinement measure. The result of six groups is indicated in Table 3.4.

Language	Number of courses	Total number of enrolment
English	21	181,412,000
Spanish	5	66,199,700
French	6	45,724,000
German	6	28,083,200
Italian	3	18,483,000
Portuguese	2	9,870,000

Table 3.3: Popular learning languages in Duolingo.

In the previous works of game refinement theory, GR -value of sophisticated games like sports and board games often fall into the range between 0.07 and 0.08 (see 2.3.4). How-

Language	B	L	GR_{Badge}
English	55.62	291.96	0.0259
Spanish	64.14	319.57	0.0250
French	72.22	346.33	0.0243
German	89	381.25	0.0243
Portuguese	68	379	0.0204
Italian	66	385	0.0200

Table 3.4: Game refinement measure of Badge in learning language groups.

ever, we noticed that the results of Badge in Duolingo much lower. Here we take English group as an example, there are 55.6 badges and 291.9 lessons on average, so the GR -value is 0.0259. That means learners, who study English in Duolingo, must complete more than 5 lessons in order to achieve only one badge in the average. Thus, the “game” is really too challenging and serious for those learners.

The GR -values of six groups fall into the range between 0.020 and 0.026. This range is lower than the appropriate range of game refinement measure in sophisticated games. The reason is that gamification applications are less fun than video or board games according to Fig.1.1.1. Moreover, as we introduced, Duolingo is a learning platform, which means that it is a serious environment and game elements are only used to increase motivation and engagement of learners, they are not used to make a course become entertaining or relaxing as fun games. Hence, this range is reasonable for gamification application and we assume the range between 0.02 and 0.03 is the appropriate range of game refinement value for gamification domain.

The Analysis of a Language Course’s Structure

With the degree of challenging which is indicated by low value of GR , the “game” in a language course only increases the motivation for advanced users or who learned with a purpose. With novice users or non-native-language learners, they give up easily their study. There is a reason why Duolingo authors have applied a milestones technique. This technique helps users split a boring and challenging learning process into several stages in order to respond learner’s efforts. For each stage, a learning process is shorter and easier to pass that helps users feel accomplished and refreshed. Like as a game, the badges in course are structured so that learners may have various “levels” of goals. Generally, the requirements of each “level” of goal get increasingly harder from completing the initial tasks until completing the course. This allows learners to learn and practice skills.

In Duolingo, users are accepted to skip some milestones by doing a test. Therefore, in the beginning, they are able to start at their suitable milestone. Moreover, the users, who are in the high level of milestone, has passed all lessons and achieved all badges from the previous milestone. For example, the users, who are in the third milestone, have passed 95 lessons and obtained 22 badges in EFSS according to data in Table 3.5.

The average number of badges and lessons are accumulated from the start milestone s to the considered milestone k (see Eq. 3.2 and Eq. 3.3). Finally, we have a formula for calculating game refinement measure in a considered milestone which is shown as Eq. 3.4.

$$B_k = \sum_{i=s}^k B \quad (3.2)$$

$$L_k = \sum_{i=s}^k L \quad (3.3)$$

$$GR_k = \frac{\sqrt{B_k}}{L_k} \quad (3.4)$$

We conduct an analysis by focusing on a course structure. Experiment’s data is collected from the most popular language course “English for Spanish speaker” (EFSS). We choose this course because it is the best course in Duolingo which attracted 142 million users. Moreover, this course has been developed for a long time and its course structure is improved during that period. Therefore, it is the best candidate for analyzing. We assume that each milestone in a course is a sub-game. Next, we calculate GR -value in each sub-game in a course by using Eq. 3.4. The collected data and GR -value in each milestone of EFSS course is expressed in Table 3.5 and Table 3.6.

Milestone	B	L
1	10	39
2	12	56
3	14	76
4	13	81
5	15	69

Table 3.5: Number of badges and lessons in EFSS.

Milestone k	B_k	L_k	GR_k
1	10	39	0.081
2	22	95	0.049
3	36	171	0.035
4	49	252	0.028
5	64	321	0.025

Table 3.6: GR -values in milestones of EFSS.

According to Table 3.6, the GR -value of each milestone shows that the milestone is designed for various types of learners. For instance, in the first milestone, GR -value is 0.081. This value is higher than the results of the sophisticated sports and board games, which

implies that the “game” in the first milestone is so exciting and attractive for beginners. The increase in the requirement is to give more challenging and exciting to learners since their skill gets better at every milestone. Therefore, GR -value decrease and maintain in two last for adapting the advancement of users. Although learners must learn a ton of knowledge actually, the milestone technique creates “rest stops” for their “learning journey” that make they reduce a sock compared to learning all lessons in one time. Therefore, a course structure helps users engage with their learning process and the game refinement trend is good in indicating the effect of a course structure.

Next, we do a brief experiment for expressing different effects which a course structure has caused on different users. In general, users usually start from the first milestone and accomplish the final goal. However, as we have stated, the advanced users can start from the second or third milestone by ignoring the previous ones. In that case, we also calculate the GR -value for those users and depict them in Fig. 3.3. We noticed that the GR -value goes down sharply in all cases and maintaining the low value after the 4th milestone. According to this figure, an advanced user who starts from the high level of milestone would feel less interesting than who start at the low level of milestone in the beginning, but they enjoy more in the end of the course.

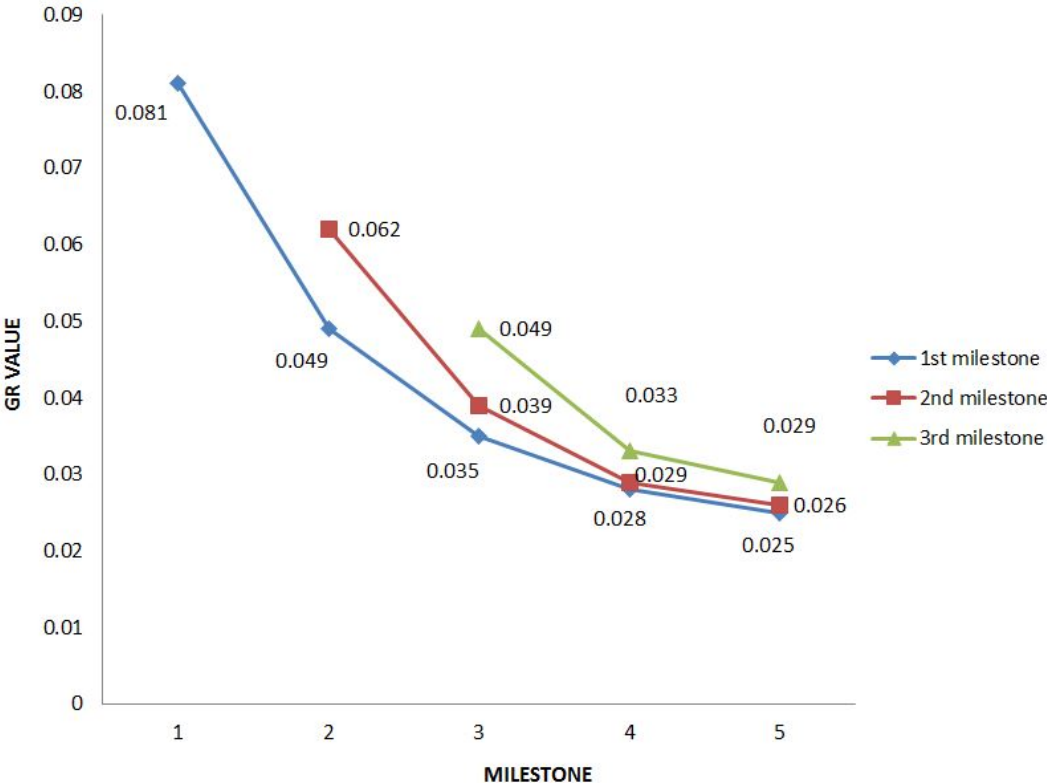


Figure 3.3: Game refinement value trends when starting from different milestones.

3.3.4 Conclusion

In this study, we have focused on the game element Badge, which has been used in the skill-tree, to analyze its attractiveness. Badge is an important game element, which is used to boost the motivation of learners when they aim to obtain a new knowledge. For the experiment, we have collected data and do an analysis on the popular learning language groups. The result has shown that the game refinement value of those cases are lower than sophisticated games. This result is reasonable because Duolingo is a serious environment which means the game elements are not used for making the entertainment environment as fun games. Moreover, the range between 0.02 and 0.03, which is much lower than an appropriate range of sophisticated games, is proposed as suitable range for GR -values in gamification applications.

Our second analysis focus on the course structure of a language course. A milestone technique was highlighted in this study. We assumed that every milestone in a course is a sub-game and figured out GR -value for them. After analyzed, the game refinement trend in the considered course indicated that the increase of challenge in each milestone aims at adapting the advancement of learners' skill. This trend also express the effect of the course structure that making rest stops for a long learning journey. Moreover, by doing a brief experiment on a course structure, we discovered that Duolingo is enjoyable for new comers who start from the first milestone, however, less enjoyable for advanced users who start at the second or the third milestone.

3.4 The Effect of a Course Structure on the Entertainment in Language Courses

3.4.1 Introduction

In the previous study, we have analyzed the attractiveness of the game element Badge in a language course of Duolingo. Moreover, we have also highlighted the milestone technique which is used to create a structure for a language course. The effect of a course structure is indicated obviously by making some checkpoints for a long and boring learning road. Based on game refinement measure, we have expressed the attractiveness of game element Badge through each milestone. The decrease of game refinement value through each milestone shows the increase of challenge in a language course which tries to adapt the advancement of learners. Although the effect of a course structure is indicated carefully in the previous analysis, we are not sure whether different course structures make any different effect on users in Duolingo. Therefore, in this study, we aim at comparing a course structure between three popular courses in Duolingo in order to explore the difference between them. Moreover, we also conduct a comparison between a structure of Japanese course in Duolingo and this in other popular learning language platform, MindSnacks.

3.4.2 Discovering Effect of Course Structure Based on the Slope of Game Refinement Trend

In this section, we analyze and compare a structure of three popular language courses in Duolingo which are “Spanish for English speakers”, “French for English speakers”, and “Italian for English speakers”. The analysis and the comparison are conducted based on the game refinement value trend and the statistic data in those courses.

Data Collection and Comparison of Language Courses’ Structures.

Duolingo has more than 29 language courses in total. However, this does not mean we can compare any couple language course together. For instance, “Japanese for Chinese speakers” and “Japanese for English speakers” cannot be compared because Japanese is easier to learn for Chinese speaker rather than English speaker. That leads to Chinese speakers may be interesting in learning Japanese rather than English speakers and we call that is the bias of a language in learning. Therefore, for ignoring this bias, we should choose the courses, which have the same difficulty in learning, for conducting our experiment. According to language difficulty ranking for English Speaker [21], three languages: Spanish, French, and Italian have the same difficulty for English speakers. For that reason, we choose three courses: “Spanish for English speakers”, “French for English speakers”, and “Italian for English speakers” as our experiment candidates in this research. The data for measuring GR -value in the course structure of three above course is collected in the same way as in Section 3.3.3. We also reuse Eq. 3.4 for measuring game refinement value in each milestone of those courses. The results are shown as follows (see three Tables 3.7, 3.8, and 3.9).

Milestone k	B_k	L_k	GR_k
1	6	21	0.117
2	16	67	0.06
3	29	160	0.034
4	61	317	0.025

Table 3.7: GR -values in milestones of Spanish for English Speakers.

Milestone k	B_k	L_k	GR_k
1	7	31	0.085
2	18	67	0.063
3	37	154	0.039
4	78	358	0.024

Table 3.8: GR -values in milestones of French for English Speakers.

According to the results, the GR -value in the first milestone of Spanish course is 0.117. It is higher than another courses (0.085 and 0.079) but the GR -values of three courses

Milestone k	B_k	L_k	GR_k
1	8	36	0.079
2	17	76	0.054
3	38	209	0.029
4	66	405	0.020

Table 3.9: GR -values in milestones of Italian for English Speakers.

are not much different from the second milestone. The first milestone's GR -values of those courses are greater than or equal to GR -values in sophisticated games. However, those values are not maintained until the end. They decrease sharply after that and fall into the appropriate range between 0.02 – 0.03 in the last milestone. Although three course structures all include four milestones, the different divisions of learning content and number of learning material in each course can effect on the enjoyment aspect. The shorter learning material and the small division in the first milestone of Spanish course make it is more attractive than two other courses (0.117 in beginning and 0.025 in the end). That leads to a number of learners in this course is higher than two other courses (109 million users compared to 64 million users in French course and 25 million users in Italian course). Therefore, we conclude that the different way used to split learning process can make different attractiveness for learners.

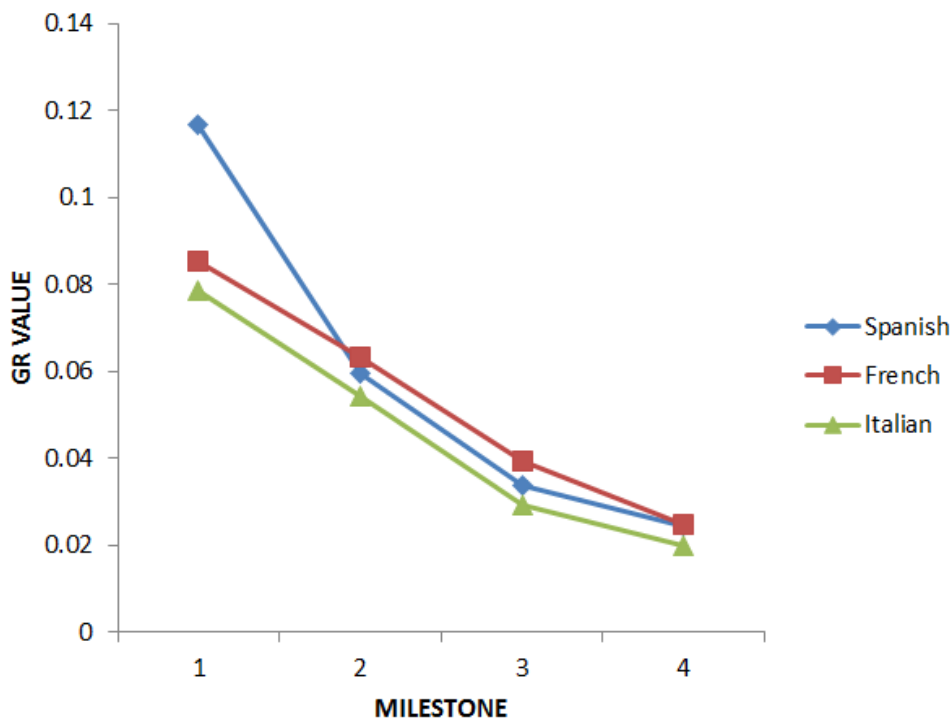


Figure 3.4: Game refinement value trends of three courses in Duolingo

Discussing on the Slope of Game Refinement Value Trend.

Although the attractiveness of Badge in Spanish source is higher than the others, the GR trend of its is less smoothly than the trends of French and Italian courses according to the comparison of the GR -value trends in Fig. 3.4. The slope from the first milestone to the second milestone in Spanish course is too sloping. The large decrease of GR -value in that period can make users feel anxiety. From the viewpoint of game designers, this is not good experience for users in game. To prove that viewpoint, we conduct an analysis on the slope of game refinement value trend. Firstly, we know that the game refinement measure presents the attractiveness of Badge. The strength of this attractiveness based on how much users focus on this game element. Therefore, the effect of game element Badge decreases if users are attracted by other game element. In Duolingo, when users complete their lessons, they not only achieve badges but also has an opportunity for obtaining or extending their winning streak. We call users, who aim at obtaining or extending their Winning Streak, is “streaking-users” and the percentage of streaking-users (PoUS) reflect the decrease of users’ engagement with the game element Badge. We do some statistic to figure out the PoUS and present it with the slope of game refinement value trend as follows.

GR Slope	GR_k	PoSU
–	0.117	9%
0.057	0.06	41%
0.026	0.034	60%
0.009	0.025	67%

Table 3.10: GR Slope and PoSU in Spanish for English Speakers.

GR Slope	GR_k	PoSU
–	0.085	10%
0.022	0.063	25%
0.024	0.039	45%
0.015	0.024	67%

Table 3.11: GR Slope and PoSU in French for English Speakers.

GR Slope	GR_k	PoSU
–	0.079	14%
0.025	0.054	34%
0.025	0.029	51%
0.009	0.020	67%

Table 3.12: GR Slope and PoSU in Italian for English Speakers.

As reported by the results in Table 3.10, Table 3.11 and Table 3.12, if the slope is more sloping, users are easier attracted by other game elements. For example, the slope, which

starts from the first milestone to the second milestone in Spanish course, is too slopping (GR slope value is 0.057). That leads to there is more than 40% users who are attracted by the game element Winning Streaks in the second milestone of this course. Whereas, there is only 25% and 34% of those in French course and Italian course correspondingly. Furthermore, the PoSU in a milestone is also effected by *GR*-values. For example, the *GR*-value in the first milestone of Italian course is lowest within three considered courses, therefore, its PoSU is highest in this milestone. While in French course, the *GR*-value in the first milestone is high enough to attract the beginner, but its PoSU in this milestone does not much different to Spanish course. However, the GR trend of French course is more stable than two other courses. That is a reason why the PoSU of this course is still lower than 50% in the third milestone while those in Italian and Spanish course are over that level. That means users in French course engage with the game element Badge for a longer period rather than two other courses.

The results also pointed out that the increase of PoUS through each milestone is related to the decrease of *GR*-value. The percentage of streaking-users in those courses reaches at 67% when *GR*-value fall into the appropriate range. That means when the *GR*-value is low, the attractiveness of the considered game element is also low. That leads to users is no longer engage with those game element. Therefore, the game refinement measure works well in presenting the attractiveness of the game element in the educational environment.

3.4.3 Comparative study: A Case Study in Analyzing Gamification between MindSnacks and Duolingo

According to the comparison between three language course structures in previous studies, we conclude that different structure can effect on the attractiveness. In order to argue this conclusion, in this section, we conduct other comparison between Duolingo and the popular learning language application MindSnacks which is evaluated by users as more interesting than Duonlingo for beginners.

Game Refinement Measure in MindSnacks

In MindSnacks, there are different mini games which are used to learn all lessons. Users are first given three available mini games to start. They can use those mini game for learning each lesson in MindSnacks. Each mini game has different game play and focuses on different area in language learning. In order to unlock other mini games, users must reach the required number of levels. That means, level in MindSnacks is used as an achievement for a learning purpose and users must level up for obtaining a new mini game. However, for level upping, users have to complete several learning tasks by playing mini games. Thus, the game concept in MindSnacks can be considered as completing several number of tasks in order to level up and unlock the mini games.

As we stated, each mini game focus on each area in language learning so unlocking a new mini game is passing a new stage in learning journey. This concept is similar to the

milestone technique in Duolingo. Therefore, unlocking a new mini game is considered as a milestone and the unlocking process is considered as a course structure. Following those considerations, we represent Lv_k for the average number of required level for unlocking the mini game (milestone) k and T_k as the average number of tasks. The GR -value in each milestone of MindSnacks is shown as Eq. 3.5 which is derived from Eq. 2.8.

$$GR_k = \frac{\sqrt{Lv_k}}{T_k} \quad (3.5)$$

Experiment and Discussion

For ignoring the bias in learning language mentioned in the previous section, data collection and analysis has been conducted on the same language course which is Japanese course. As in MindSnacks Japanese, eight mini games are provided in total. We have recorded such data which is the average number of required levels and the average number of required tasks for unlocking the mini games in this course. From the collected data, the GR -value of each milestone is measured using the Eq. 3.5 and shown in Table 3.13. While in Duolingo, we have followed the same method of data collection as in Section 3.3.3. We also reuse Eq. 3.4 for calculating the GR -value of each milestone is measured in Duolingo. The collected data and result are indicated in Table 3.14.

Mini games	Lv_k	T_k	GR_k
Totem	3	5	0.35
Bloon	6	17	0.14
Stacks	10	37	0.09
Bubbler	15	85	0.05
Dam Builder	21	141	0.03

Table 3.13: GR -values in milestones of MindSnacks Japanese

Milestone	B_k	L_k	GR_k
1	7	34	0.08
2	15	66	0.06
3	27	130	0.04
4	40	185	0.03

Table 3.14: GR -values in milestones of Duolingo Japanese

The previous section shows that the GR -values of Duolingo language courses fall into the range between 0.020 and 0.03 which is considered reasonable for the serious environment game-like setting. The similar tendency happens in MindSnacks and Duolingo Japanese that the GR -value for each milestone decreases gradually from 0.35 to 0.03 and 0.08 to 0.03 respectively (refer to Fig. 3.5). There is a big gap between GR -value at the first milestone of both platforms due to the big difference of the average number of tasks or

lessons needed to complete the first milestone. This difference in value indicates that MindSnacks Japanese is more entertaining and attractive to be used for beginners rather than Duolingo. However, this big difference in the first milestone is not maintained until the end. The GR -value of MindSnacks and Duolingo starting from the first milestone

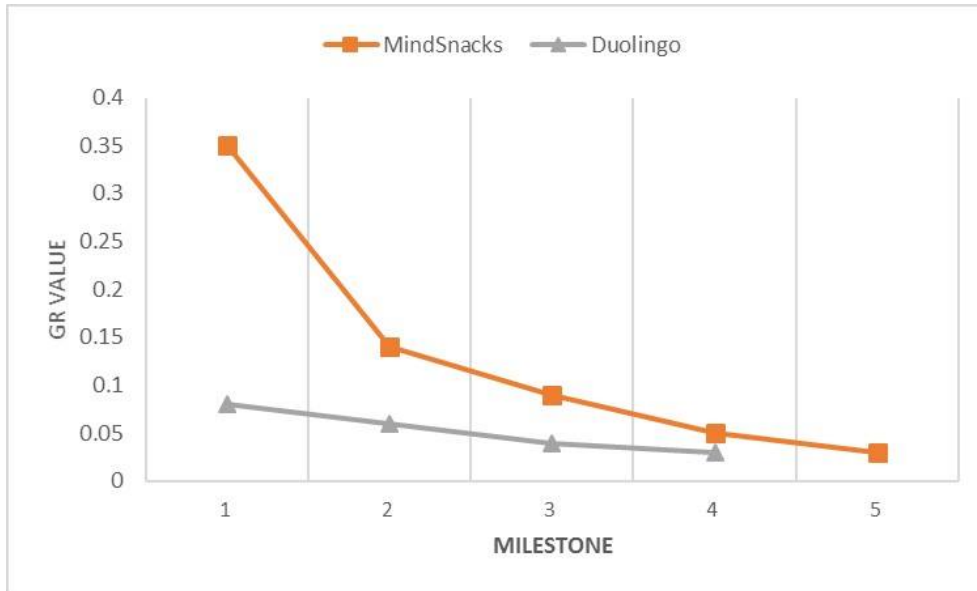


Figure 3.5: Game refinement value trends of MindSnacks and Duolingo

started to decline and result in a small difference of GR -value as user progress through the milestone. At the end of each milestone, MindSnacks and Duolingo result in the same GR -value which is 0.03 at the fifth and fourth milestone respectively. In the terms of milestones, MindSnacks has one extra milestone compared to Duolingo. Considering the pattern where the GR -value of MindSnacks turns into 0.03 at the fifth milestone whereas Duolingo at the fourth milestone, this could be seen as the longer it takes to reach a milestone in learning, the higher the excitement and motivation of users learning. A longer milestone gives an effect on the users' feeling where more milestones give more sense of achievement. The number of accumulated tasks at the end of each milestone gives a difference in value where MindSnacks has a lower total number of tasks than Duolingo, which means that the lesser number of tasks to reach a milestone, the higher the enjoyment and motivation to finish the lessons. According to the comparisons and differences of GR -value, MindSnacks is more interesting than Duolingo at the beginning. Some factors that may cause this result is due to its courses structure as summarized in Table 3.15.

3.4.4 Conclusion

In this study, we have compared the course structure of three courses for figuring out the effects of its. The data collection way and GR -value measuring formula are reused from the previous studies. After the experiment and comparison have been conducted,

MindSnacks	Comparison Factors	Duolingo
141 tasks	Number of accumulated tasks/lessons at the end of milestone	185 lessons
Flexible; select lessons based on user preference	Learning Structure	Step by step learning; required to unlock skills before proceeding to other lessons
Game-like approach	Questions Structure	Quiz-like approach
No; to access all lessons required payment	Free platform	Yes

Table 3.15: Comparison factors between MindSnacks and Duolingo

the results show that the division and number of learning material in a structure have an influence on the attractiveness of this course. Furthermore, based on the slope of game refinement in the course structure, we have figured out that if the slope is more sloping, users are easier attracted by other game elements. The results also indicated that the structure of French course is better than the structure of two other courses. Moreover, the decrease of game refinement value through each milestone can express the decrease of user engagement to the game element. Therefore, from those point of views, we concluded that the game refinement measure works well in presenting the attractiveness of considered game element.

A comparative study between MindSnacks and Duolingo shows the application and game refinement measure reliability in the serious game-like environment. By using the derivation of the game progress model of an educational game, we have quantified the attractiveness and entertainment impact of game where the range of GR -values resulted in the appropriate zone of serious games. As the result, the differences of GR -values between these applications contribute to the significance in interpreting learners' enjoyment points. The comparison pattern of GR -value shows that MindSnacks' learners enjoy the use of gamified learning approach rather than Duolingo. From this study, we figure out that more aspect of gamified learning is discovered to foresee the use and impact of game elements in an educational learning environment as well as to acknowledge the need for additional research in the serious environment.

3.5 Analyzing Winning Streak's Effects in Language Course of Duolingo

This study explores the effects of the game element Winning Streak on users' motivation and engagement in Duolingo's language course. The game element Winning Streak has

been used in sport and video games to describe a consecutive number of successful actions and increase players' attention to complete their goal. Similarly, in gamified education system, Winning Streak is employed as a game element to boost up motivation of learners. By applying game refinement theory as an assessment method, enjoyment of two user groups in Duolingo is measured to compare in this study.

3.5.1 Introduction

Duolingo's authors have used many game element for helping the learning process of users become more entertainment. Although Duolingo users focus mainly on the game element Badge, no one can deny the helpful effects of others game element on learning process of users such as Winning Streak, Leader Board. Therefore, we should analyze other game element for more understanding this system. In this section, we choose the game element Winning Streak as our next target. This game element is chosen for two reasons. Firstly, a winning streak can be obtained by completing a learning lesson. That means when users complete their learning lesson, they not only obtain a badge but also has a chance to extend the length of their winning streak. The second reason is the game element Winning Streak has been used in many sophisticated games such as basketball, chess, soccer, or DOTA2 in order to increase players motivation and performance. In Duolingo, the game element Winning Streak looks like also effect on learning process of users. According to Table 3.17, we see the percentage of streaking-users increases when the attractiveness of game element Badge decreases. Hence, in this study, we aim at analyzing the effects of the game element Winning Streak in Duolingo in order to clarify its effects as well as its contribution in entertainment aspect of Duolingo. In this research, we conduct two analyses. The first analysis aim at figuring out the individual attractiveness of Winning Streak in a language course. While the second aim at discovering its contribution in entertainment aspect by analyzing the combination between Winning Streak and the game element Badge.

3.5.2 Winning Streak Attractiveness

What is a Winning Streak?

The term Winning Streak was initially used in sports [35]. It refers to a consecutive number of games won, which begins from the third consecutive victory. Winning Streak is held not only by an individual, as in tennis, but also by a team. For example, we mention basketball, soccer and hockey. In basketball, a hot hand was used to describe a basketball player who had been very successful in scoring over a short period. It was believed that players who make a shot are more likely to hit the next shot than players who miss a shot [3]. Hence, the players, who obtained a winning chain, always keep their streak continue. In other words, Winning Streak effects directly on players' attention in order to complete their goal for increasing their performance. Furthermore, Winning Streak is also used in video games. For example, in DOTA2, a killing streak is used to increase players' attention to conduct a battle with opponent players. Corresponding to

the number of kills, some titles are awarded to players [8], such as Killing Spree (3 kills), Dominating (4 kills), Mega Kill (5 kills), etc., in order to encourage users to repeat their actions. There are many famous video games using Winning Streak in the game play, such as Call of Duty, Cross Fire and League of Legends. In Duolingo, Winning Streak has been used as a measurement of how consistently learners study. A winning streak of users starts at zero and increases by one as long as they complete to study their required number of lessons. The winning streak will be reset to zero if the number of studied lessons is missed. However, the learners can keep their streak by using “streak freeze” which is sold in Duolingo shop.

Game Refinement Measure for Winning Streak

We figure out the game progress model of the game element Winning Streak based on extending the streak’s length of diligent learners. A winning streak can be obtained by complete lessons in order to reach the learning goal. Although it is easy to obtain, extending the length of winning streak is more difficult. For example, if we want to extend our winning streak’s length to ten, we must pass the learning goal in ten consecutive days. That is a reason why the longer winning is more precious. Therefore, the main game process can be given by two factors which are the number of learning consecutive day and total number of days within a year. As the total number of days in year is 365 and the number of learning consecutive day is the length of a winning streak. Thus, the game refinement measure for game element Winning Streak in each milestone is given by Eq. 3.6 where W_k is the average length of a winning streak in a milestone k .

$$GR_k = \frac{\sqrt{W_k}}{365} \quad (3.6)$$

Experiment and Discussion

Each language course in Duolingo has different learning material and structure. That means each course is an individual learning environment. Therefore, to see the effects of a winning streak obviously, we selected 2000 users, who enrolled the most popular course English for Spanish Speakers (EFSS), to collect data. The criterion, which is used to select candidates, is that they must be active on the system within one week before the data collecting date. Because some information fields of users should be kept in secret, so we only collected public information of users. The collected information, which is indicated in user’s profile page of Duolingo, includes a number of skills, length of a winning streak and the languages course. After the data was collected, we use Eq. 3.6 for measuring attractiveness of game element Winning Streak. The result is shown in Table 3.16.

According to Table 3.16, the game refinement value of Winning Streak increases through each milestone. It starts from 0.004, increases almost double after that and reaches at value 0.021 in the last milestone. Similar to GR -value of Badge, the GR -value of Winning Streak also falls into the appropriate range in the end. This is one more evidence which proves that the range between 0.02 – 0.03 is should be considered as the appropriate range for gamification domain. Whereas obtaining a badge is used as a gaming activity

Milestone	W_k	GR_k
1	2.2	0.004
2	6.4	0.007
3	21.9	0.013
4	47.8	0.019
5	59.1	0.021

Table 3.16: GR -values for Winning Streak in milestones of EFSS

for attracting beginner to acquire new knowledge, extending the length of a winning streak is considered as enhancing normal learning activity to serious game activity. According to Fig 3.6, game refinement value of Badge starts from the appropriate range of sophisticated games which means users are attracted to learn as play a fun game in the beginning. However, the GR -value decreases and maintains at appropriate range of gamification domain after that. Obtaining badges is an interesting activity but do it frequently will become boring. That explains why the attractiveness of Badge decrease after a period of time. On contrary, learning language is actually an normal activity. For making it become interesting, the author of Duolingo have used Winning Streak as a reward for users who are diligent. The length of winning streak also present user's efforts. Therefore, the longer winning streak is more precious. This streak is shown on users profile and it reflect the ability of users in learning. In consequence, the experiment indicated that the winning streak is really effect on users learning process. It is able to help Duolingo improve entertainment aspect when attractiveness of Badge decrease. To clarify that point of view we will do an analysis on the contribution of Winning Streak in the improvement of entertainment in the course EFSS. The detail of this analysis is presented in the next section.

3.5.3 The contribution of Winning Streak in the Improvement of Entertainment

Attractiveness of Badge and Winning Streak

The previous study (Section 3.3.3) expressed the use of milestone technique in Duolingo. This technique used to split a big task into smaller tasks in order to help users complete their works easier. Therefore, we consider each milestone in a language course as a sub game and measure the game refinement value for those sub games in this course. In general, the normal users do not study consecutively and they only pay their attention to getting badges, so that the game refinement value of each milestone is measured by Eq. 3.4. Actually, there is not only Badge considered as an achievement. A winning streak is an extra reward for users, who always keep their study in progress, which means that they have to finish learning a consistent number of lessons every day. We call the user, who obtained a winning streak, a streaking-user while others are normal-users. According to the percentage of streaking-users in each milestone, the importance of obtaining streaks is different in different level of milestone. Therefore, we apply Eq. 3.7 in order to measure

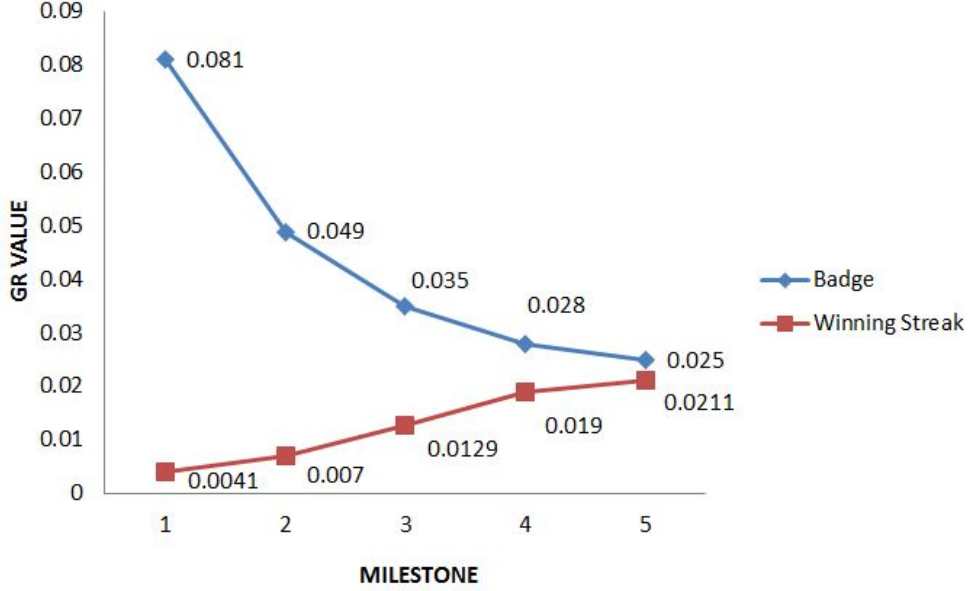


Figure 3.6: Game refinement value trends of Winning Streak and Badge.

attractiveness of a game element Winning Streak in a milestone k . In Eq. 3.7, B_k and L_k denote the average number of obtained badges and the average number of learned lessons, where i_k stand for the importance value of a winning streak in a milestone k , and W_k stands for the average length of winning streak in this milestone.

$$GR_k = \frac{\sqrt{B_k + i_k W_k}}{L_k} \quad (3.7)$$

The importance value of winning streak expresses the strength of attractiveness caused by the streak in each milestone. Therefore, we calculate it based on the percentage of streaking-users. At high levels, new lessons are not easy for users to be obtained, so that a winning streak is very important for users' engagement at those levels. Hence, the importance value of winning streak accumulated calculates by Eq. 3.8, where i_k denotes the importance value of streaks and $PoSU_k$ stands for the percentage of streaking-users in the milestone k in decimal. Although in some case a winning streak is not important, nobody can deny its attractiveness. Therefore the parameter i_k should start from 1.

$$i_k = i_{k-1} + PoSU_k \quad (3.8)$$

Data Collection and Analysis

In this study, we also use a data set which is collected from EFSS course in Section 3.5.2. Based on collected data, we analyze the statistics to figure out the percentage of streaking-users (PoSU) in each milestone. The statistical results and the importance values of

Milestone k	$PoSU$	i_k
1	15%	1.15
2	23%	1.38
3	27%	1.65
4	30%	1.95
5	57%	2.52

Table 3.17: Percentage of streaking-users in each milestone.

winning streak are calculated as shown in Table 3.17. After that, we measured two kinds of game refinement values in each milestone in order to evaluate the effects of winning streak by comparing them. The first group consists of streaking-users and another group contains normal users. The game refinement values GR_k of the first group are measured by Eq. 3.7 while the second group calculated by Eq. 3.4. Consequently, the results are presented in Table 3.18.

Milestone k	B_k	L_k	i_k	W_k	GR_k	$GR_k(\text{normal})$	Difference	Increase(%)
1	10	39	1.15	2.2	0.09	0.081	0.009	11.11
2	22	95	1.38	6.4	0.06	0.049	0.009	18.37
3	36	171	1.65	21.9	0.05	0.035	0.015	42.86
4	49	252	1.95	47.8	0.047	0.028	0.019	67.86
5	64	321	2.52	59.1	0.045	0.025	0.02	80

Table 3.18: Game refinement values in detail of each milestone in EFSS.

Table 3.18 shows that the difference of GR -values between two user groups increases through each milestone. This implies that a winning streak can increase attractiveness in each milestone of EFSS at least by 0.009 (see the case $k=1$ and $k=2$) for streaking-users. The difference of the fifth milestone is 0.02, while the difference between the first and the second milestone are trivial. Let us consider the trend of GR -values for normal-users. The decrease of game refinement values indicates that the game challenge increases to adapt the advancement of learners' skill. However, we are not certain that the increase can adapt all learners' advancement in a course. In high levels, users are required to have a certain level of knowledge in order to pass all lessons in each skill. Therefore, it is more difficult to obtain badges. Someone, who has not enough knowledge to pass lessons at high levels, should practice more to improve their skills. Nonetheless, a badge is not given by reviewing lessons. That is a reason why a winning streak is only the element that can improve users' motivation in this case. For this reason, the winning streak is necessary at those levels. It means that a winning streak can improve attractiveness of the game when the challenge increases. The winning streak becomes more significant for advanced users, who are at high level of milestone, rather than beginners. According to the streaking-users percentage which increases by each milestone (see Table 3.17), over 56% users in the fifth milestone keep their study progress by obtaining a winning streak, whereas only 14.82% users do that in the first milestone. Moreover, considering

the difference of GR -values between two user groups as the effects caused by a winning streak on streaking-users, the winning streak can increase attractiveness up to 80% in the last milestone, but it is less than 20% in two first milestones. Finally, we will prove our analysis which is conducted based on game refinement value by using the real collected data. We collect data of 2000 users again after 5 months from the first time. Those data, which is shown in Table 3.19, express that streaking-users have higher performance than normal users. The average of obtained skill of streaking-users is 15.81 while this of normal users is only 9.11. Therefore, a winning streak help increase users' motivation for completing their goal. Moreover, based on the given up percentage, Winning Streak also help increase users engagement because the given up percentage of normal users (61.7%) is much higher than the percentage of streaking-users (21.1%). Last but not least, this game element is more important for advanced users than beginner as our analysis as well (almost 50% streaking-users in the last milestone).

Compared Factors	Streaking users	Normal users
Average number of obtained skills	15.81	9.11
Given up percentage	21.1%	61.7%
Active users distribution in 5 milestones	Milestone 1: 3.6%	Milestone 1: 96.4%
	Milestone 2: 7.9%	Milestone 2: 92.1%
	Milestone 3: 8.2%	Milestone 3: 91.8%
	Milestone 4: 20.1%	Milestone 4: 79.9%
	Milestone 5: 48.9%	Milestone 5: 51.1%

Table 3.19: Real statistic data in EFSS.

3.5.4 Conclusion

This study aims at exploring the effects of the game element Winning Streak on the learning process of users in Duolingo. In this research, we first figure out the attractiveness of Winning Streak individually. Then, we discover its contribution in the improvement of entertainment aspect. For the experiment, we collected data of 2000 users to measure a game refinement value of each milestone in the most popular language course “English for Spanish speakers”. Although we should consider other language courses, the first result is enough to indicate that the use of the game element Winning Streak in Duolingo caused the effects on motivation of users as in video games and sports games. The result of the first consideration has shown that Winning Streak helps users enhance their normal learning activity to serious game activity. That means users is interested their learning activity with Winning Streak. The longer winning streak is more precious for users in Duolingo. After compared the game refinement value trend of Winning Streak and Badge, we recognized that Winning Streak helps Duolingo improve its enjoyment when the attractiveness of Badge decreases. Therefore, the second analysis was conducted in order to clarify this view point. According to the difference of GR -values between two kinds of

users, the results show that the streaking-users are more attracted rather than the normal users. Additionally, by comparing the increase ratios of attractiveness and the streaking-users percentages between milestones, the results also expressed that a winning streak is more significant for advanced users rather than beginners. In conclusion, a winning streak helps to increase users attention to their learning purpose when the challenges increase. It also works well in improving motivation of users because of its attractiveness. When a winning streak reaches an appreciable length, it becomes more precious. Users do not want to lose their winning streak because they have to pay many efforts to reach that length. Therefore, users have to keep their learning process go ahead every day for winning streak lengthening. That leads to a winning streak increases its attractiveness as it lengthens. Those analyses is proven by giving a really statistic data in considered course.

3.6 Chapter Summary

This chapter presents the first application of game refinement theory in educational game domain. The game refinement measure has been calculated for the game elements Badge. By using Duolingo as our main testbed, we have figured out the GR -value of this game element. As reported by experimenting, GR -values of Badge in the popular language courses fall into the range between 0.02 to 0.03. Therefore, we have proposed this range as the appropriate range for educational game domain. This range is lower than the appropriate range in sophisticated games. Hence, the Duolingo authors have used milestone technique to split a long and boring learning process into several shorter processes which users can pass easily. That makes Duolingo become more interesting. The authors use this technique to create rest stops in the long learning journey which users can rest and review their efforts. We called the split way of learning material by using milestone technique is a course structure. Based on the the comparison of GR -value trends in different course structures, we concluded that the course structure effects on the attractiveness in the beginning of the game and the engagement of users on the game element Badge.

Also in this chapter, we have extended our idea to analysis other game element in Duolingo which is the Winning Streak. Winning Streak is used popularly in sport and video games for increasing the focus of player to complete their goal. In this research, we first figure out the attractiveness of Winning Streak individually. Then, we discover its contribution in the improvement of entertainment aspect. Different to Badge, the game refinement trend of Winning Streak increases through each milestone. In the beginning Winning Streak cannot attract users because it start at nearest 0 point of GR -values which means obtaining a winning streak at this time similar to a normal activity. However, when the attractiveness of Badge decreases, users tend to be attracted by other game element and in this case Winning Streak is the best choice. After conducted the analysis on the data of 2000 users, we also figured out the contribution of Winning Streak to the improvement of entertainment aspect when the attractiveness of Badge decrease. Our analysis also be confronted with the real statistic data.

Until now we have do some analyses on game elements in educational game Duolingo for understanding their effect obviously. By using game refinement measure as the attractiveness measurement for game element, we have figured out and interpreted the effect of game element successfully. Therefore, the game refinement value can be a good measurement to help game designer can decide their strategy in using game element in gamification platform.

Chapter 4

Design and Evaluation of Cybersecurity Awareness Training Game

Serious games are becoming more popular because they provide an opportunity for learning in a natural environment. Although many concepts included in cybersecurity awareness training are universal, many forms of training fail because they are based on rote learning and do not require users to think about security concepts. The main objective of this study is to design a candidate cybersecurity awareness training tool which provides an environment for helping users to understand security concepts while playing a game. To reach our goal, we applied a newly developed model, which is Activity Theory-based Model of Serious Games (ATMSG), to design our own game. According to the design, we implemented a game demo and assessed its game play. The results indicated that the story game play can help players improve their understanding of cybersecurity problems and resolutions.

4.1 Introduction

Nowadays, the whole world is facing severe challenges posed by hackers, spammers and a large pool of attackers who are motivated by a variety of reasons. To prevent attacks from high-tech criminals, some organizations use a cybersecurity training tool to enhance its information assurance posture. However, creating an engaging training tool, which holds trainees' attention sufficiently long to impart awareness, is a considerable challenge. Moreover, many forms of training fail the studied concepts because the training is by rote learning and does not require users to think about and apply. Generally, to increase motivation and engagement of users, many trainers has designed their training tools as a serious game. A serious game is the combination between game learning sciences and digital entertainment. Similar to simulations, serious games present a virtual reality of varying fidelity that allows learners to explore, experiment, or simply engage in learning. Therefore, our research aims at designing a training tool, which helps users enhance their

cybersecurity knowledge, by using a serious game design model. Moreover, we also try to find a suitable game play which helps users understand the knowledge obviously. In this study, we use the Activity Theory-based Model of Serious Games (ATMSG) [1] and a story game play to design the training game. The details are shown in the next sections.

4.2 Theoretical Background

The learning effects of serious games in studies across educational contexts are inconclusive [37]. One of the recommendations is to ensure that game objectives and learning objectives correspond. It is really difficult to determine whether or not a game supports the learning of students because a learning objective and outcomes are unclear. There exist several models such as RETAIN [14], DODDEL [23], and the 5/10 method [18] which support to design serious games. However, in this study, we use ATMSG [1] to implement our game design idea. The ATMSG provides a comprehensive way to investigate, in detail, how a serious game is structured, and uses activity theory as the theoretical background.

In ATMSG, educational serious games are seen as used in the context of four activities: the gaming activity, the learning activity, the intrinsic instructional activity and the extrinsic instructional activity. There is a four-steps-approach that progressively guides the user in applying the ATMSG to the design or analysis of educational serious games. These steps take the user from a high-level understanding of the activities to the concrete components which implement those activities. The user identifies game components with the help of the taxonomy of serious game components. The details of the four-steps-approach are shown in [1].

In comparison to other models, ATMSG offers a more precise model for the analysis of the educational and gaming aspects of a game, allowing users to perform a more exhaustive decomposition of components as the game unfolds, and to link these components to the overall learning objectives [1]. Moreover, the design using ATMSG is expressed as tables and charts. That facilitates game developers to understand the designer's idea. That is a reason why we choose ATMSG. In our design, we use a game play in story games as our main game play. We call it a "story game play". The story game play uses storytelling technique to lead players to experience game events. Hence, we apply this game play into our game to support users in clearly understanding the given problem and its solution.

4.3 Applying ATMSG to Design a Cybersecurity Awareness Game

Firstly, we introduce our game design idea. Our game design idea focuses on making a cybersecurity training game for users to learn and understand obviously cybersecurity

concepts. We aim at giving them some reality situations of security attack. For example, “How to recognize phishing emails?”, “How to protect your information at public locations?”. In those situations, users will control their character and find the solution for given problems. We will lead users through a story step by step and they must make choices, which reflect their character’s behavior. After each situation, a system will evaluate users’ answer and gives them the right answer with its explanation. Following this idea, we design a game and describe it in detail by using the four-steps-approach of ATMSG. Critical points of each step are shown in the next discussion.

Step 1 - Describe the activities: In this step, we highlight the main aspects of activities to understand this game easily. Moreover, to make it convenient to the implementation, test and evaluation, we choose the topic “Cybersecurity awareness training in a university”. It not only makes convenient for us but also players (university students) to be familiar with given situations in the game and to keep cybersecurity concepts in mind easily. As a result, the main subject of gaming and learning activities is university students who will gain knowledge by experiencing directly real cybersecurity problems and learning to avoid and solve them. Our game demo aims at providing the training tool for students in our university (JAIST). Therefore, game resources such as images, learning content or story reflect the daily life at JAIST.

Step 2 - Represent the game sequence: In this step, we provide a diagram to present our game sequence by using UML. The game includes two main activities: “Problem Solving” and “Practice”. Both game activities have different game play. The Problem Solving will give players a new situation (story) to experience. Players must understand given situations and choose their behavior carefully to get a perfect score. If players choose a wrong answer, they cannot get a score in this event. An explanation and a right answer will be given after evaluation. While Practice gives players a quiz game, which contains questions about cybersecurity awareness. Players must answer them repeatedly and they will be punished if they choose a wrong answer. Whereas Problem Solving is commonly used to support players in understanding situations, Practice helps players remember cybersecurity concepts. Practice’s game play is easy to understand, so in our demo, we only give an example. On the other hand, Problem Solving is more complicated than Practice, and our research tries to examine the effect of story game play at understanding of users. Thus, we only focus on designing Problem Solving for our game demo.

Step 3 - Identify actions, tools and goals: In order to understand a game sequence easily, we are going to identify components that are related to each node in the game sequence. Those nodes are composed from their actions, tools and goals. We first choose the relevant components directly from the taxonomy of serious games (Table 11 in [1]), and fill them into the three layers of each activity involved (gaming, learning, intrinsic and extrinsic instruction). The extrinsic instructional activity is performed outside by the teacher or instructor in the context of the overall learning setting. Our game purpose is not to create an open learning environment for a teacher to teach anything they want.

We provide a training tool which helps players in enhancing their awareness about cybersecurity problems in a specific organization (e.g. university, company, etc.). Therefore, in our case, we do not consider an extrinsic instruction, but we simply fill nine layers in total. Almost all components were selected based on main game play in order to clearly show designer’s ideas by the game actions, tools and goals. Therefore, this step can help game developers understand the designer’s idea thoroughly.

Step 4 - Description of the implementation: In this step, we provide a more detailed description of our implementation. We explain what is being done, using which tool, and with what purpose in each block of a table. We also explain how the use of such components and characteristics support the achievement of the entertainment and/or pedagogical goals of the game.

The combination of the four steps described above provides a comprehensive view of the structure of our game, from its high-level purposes and general characteristics to its concrete implementation. In this study, we used Novelty [34] to implement our design. Novelty provides simple methods to create our own visual novel game. A visual novel game [33], also known as a story game, is an interactive game. Typically, the majority of players’ interaction is limited to clicking to keep the text, graphics and sound moving on while making narrative choices along the way. Our detailed design and demo is available at [4].



Figure 4.1: Screenshots of CSAG Demo

4.4 Evaluation and Discussion

To collect players’ feedback on the demo, we created a survey which is available at [5]. The survey requires players to evaluate the demo in multiple aspects of the game play and educational value. They must rate on a scale from 1 (worst) to 10 (best). We had 10 participants who filled out the questionnaire. Six of them are at the beginner level in cybersecurity knowledge and other participants are at intermediate level. The details

of the survey question and its results are shown in Table 4.1 and Table 4.2. The survey questions are categorized into three groups. The first group (Q1–Q6) is used to test abilities of the game demo such as clarity, game length, content, and enjoyment aspect. It helps us figure out a comfortable game setting in official implementation. The second group (Q7–Q9) aims at evaluating learning purposes and the effect of story game play on helping players understand cybersecurity concept. The last question is used to estimate how much this game improves players’ motivation.

ID	Question
Q1	Clarity on what to do
Q2	Clarity of story
Q3	How fun it is to play the game
Q4	How easy it is to play the game
Q5	Length of the game
Q6	The game provides you to new material
Q7	The game helps you to understand a situation of the problem
Q8	The game helps you to understand how to solve the problem
Q9	Improving your understanding in cybersecurity
Q10	Improving the motivation to learn more about cybersecurity

Table 4.1: Survey questions.

Level	Q1	Q2	Q3	Q4	Q5*	Q6	Q7	Q8	Q9	Q10
Beginner	8.8	7.3	6.0	9.8	6.5	7.8	8.8	8.0	8.0	6.7
Intermediate	8.5	7.8	7.3	9.3	5.5	7.0	7.8	7.5	7.0	6.8
Total	8.7	7.5	6.5	9.6	6.1	7.5	8.4	7.8	7.6	6.7

* this aspect is rated on a scale from 1 (short) to 10 (long)

Table 4.2: Survey results: average score per question.

The responses to the survey show some advantages of the game. The first advantage is that the game is very easy to play for both groups of participants (Q4, the average in total is 9.6). Besides, the game is rated highly on two aspects that are what to do (Q1, the average in total is 8.7) and the story of the game (Q2, the average in total is 7.5). However, the interest of this game is rated not so high by beginners (Q3, the average is 6.0), but it is interesting for learners who are at intermediate level (Q3, the average is 7.3). The reason is that the game provides new materials to beginners more than intermediate learners (Q6, the averages are 7.8 and 7.0). That makes the beginners feel that the game length is quite long (Q5, the average is 6.5), so they become boring when playing this game. On the contrary, the game length is suitable for intermediate learners (Q5, the average is 5.5) so they feel more comfortable than the others. According to the results of questions from 7 to 9, we have high ratings on the aspect of understanding a problem

situation, how to solve the problem and improving students' understanding (Q7–Q9, the averages in total are 8.4, 7.8 and 7.6). However, those aspects are rated by beginner higher than by intermediate learners. The reason is that the game was designed to be a training tool for awareness raising for a beginner to obtain new knowledge. Therefore, this game is more helpful for beginner than intermediate learners. It seems that the game is not good at making enjoyment, so it improves the motivation of users not so well (Q10, the average in total is 6.7). In conclusion, the advantage is that the game can help the students improve their understanding of cybersecurity awareness. By giving a story, the game aims to lead players to understand cybersecurity problems and resolutions, even though it teaches an intermediate student nothing new in the material.

4.5 Designing Learning Structure for Improving the Entertainment Aspect

As we stated in the previous session, the game demo was evaluated by players as not interesting. The reasons are the game is long and learning material is too much. In the game demo, we have provided only one story from the beginning to the end with many problem events. Due to the story does not finish in the end of a problem event so that the badge, which is collected when players passed the event, is considered as an item in game but not an achievement. That leads to players only feel successful only one time in the end of the game. For improving this issue, we apply the technique which has been used in Duolingo in order to create our learning structure. We also used game refinement measure for quantifying the attractiveness of our learning structure.

Our learning content are shown in Table 4.3. Badges are given to players after they pass all problem events in skill. In previous demo, players have to pass 20 problem events but get only one achievement after all. While, in the new design, we split a long story into 4 shorter stories. In each story players will experience some problem events of a specific security topic. After each short story, players will obtain a badge as the reward for their efforts. The game refinement measure of this game can be calculated by the Eq. 4.1 where A is the average number of achievements and E is the average number of problem events. The GR -values of new version and the old version are indicated in Table 4.4. According to that table, we can see GR -value of previous version is only 0.04. It is lower than GR -value of sophisticated games so it is not enough to attract a beginner. That is a reason why players, who is in beginner level, rated that this game is not interesting. By creating the learning structure, we successfully increase GR -value to 0.08. This degree is suitable for a beginner rather than the previous one. Although the learning structure will change when the number of learning material increase, we redesign or adjust this structure by using the Eq. 4.1.

$$GR = \frac{\sqrt{A}}{E} \quad (4.1)$$

Skill	Sub-topic	Problem event
Computer Using	Ethical Use	Installing new OS
		Installing necessary softwares
		Ethical use
	Physical Security	Computer Protection
		Sensitive Information Protection
	Account and Password Management	Low Security Account Credentials
	Password Strengthen	
Email Using	Social Engineer Inboxes	Bad traffic to the Internet
	Phishing Emails	Scams Email
		Trojan attachment
		Spoofing an email address
Devices Security	Company Devices	Using Company Devices Policy
	Secure in Public Location	Public network access
		Email sending in public
		Strangers helping
Web Browser Security	Social Network	Public Information
		Friend Request
		Post Management
	Web Browsing	Cookies
		Popups

Table 4.3: Cybersecurity awareness learning content.

Version	B_k	L_k	GR
Old	1	20	0.04
New	4	20	0.08

Table 4.4: GR -values in CSAG

4.6 Chapter Summary

In this study, we presented the design of a security awareness training tool. By using the ATMSG paradigm, which offers a more precise model for the design of the educational and gaming aspects of a game, we present our game idea in detail. Moreover, the model ensures that game objectives and learning objectives correspond. The four-steps-approach helps game developers follow the idea of designer easily. To evaluate the suitability of the game structure and a story game play, we have built a game demo which is implemented by using the Novelty software. After the game was played and evaluated by university students, the result indicated that the game was rated well on the aspect of understanding cybersecurity problems and solutions. Therefore, the game which we developed by following this design help us avoid the rote learning of users when using the training tool. However, to improve the enjoyment of this game and the learners' motivation, we should add more game actions and game elements.

Although the game can help users in understanding the cybersecurity concept, it is not good at making enjoyment for players because it is simple to play by clicking and reading. Moreover the balance between tasks and achievements in a game demo is unbalancing. That also leads to the boring in the game. Among those problem issues, the balance between tasks and achievements is the most important because it is related to the learning material. Therefore, we have improved the entertainment aspect by focusing in this point. For the improvement, we have used the gamification technique to design the learning structure. After create the structure we used game refinement measurement for adjusting its. The new structure of this game has the game refinement value falls into appropriate range of sophisticated games.

Chapter 5

Conclusion

The aim of this thesis is understanding the effects of game elements in educational environment obviously for using them effectively. That is why the problem statement of this research is ‘How to use a specific game element effectively?’. To achieve our goal, we aimed at analyzing the most popular educational gamified platforms for understanding the effects of used game elements and their characteristic. We also aimed at giving a specific analysis method by presenting many case studies. Our analyses were carried on by using game refinement theory idea. The game refinement theory shows a particular way to quantify the attractiveness by providing the common measurement which is the game refinement measure.

First of all, we gave the fundamental idea of game refinement theory and its previous studies in Chapter 2. Next, we applied this idea into gamification domain, specifically educational games, and figured out the general game progress model for educational games based on the main game action. After that, we used this general model to conduct our experiment in many case studies. For the first approach, we have applied game refinement theory idea to quantify the attractiveness of game element Badge and discussed based on GR -value. The result shows that the GR -value of all popular language courses in Duolingo fall into the range between 0.02 – 0.03. We also recommended this range as the appropriate range in gamification domain. In this research, we highlighted the milestone technique as well. This technique has been used by Duolingo authors for splitting the long and boring learning journey into several shorter ones. Creating milestone is considered as making a rest stops where users can take a rest and review their efforts after passed a long learning path. The splitting way on the learning material or process of users is called a ‘course structure’. After deep analyzed, we figured out that the course structure can effect on the entertainment aspect in Duolingo. Therefore, we compared the course structures between three language courses in Duolingo as well as between Duolingo and MindSnacks to see this effect. The comparative results showed that the different course structure can make different effect on the attractiveness for beginners and the engagement of users with the considered game elements.

The second application were conducted by analyzing the effects of the game element

Winning Streak. This game element has been used popularly in sports games and video games. In Duolingo, Winning Streak has been used as a measurement of how consistently learners using Duolingo for their studying. To reach our goal, we conducted two analyses. In the first one, we analyzed the attractiveness of Winning Streak particularly. The result of this analysis showed that Winning Streak helps Duolingo users enhance their normal learning activity to serious game activity. The length of winning streak also presents users effort. Therefore, the longer winning streak is more precious. Different to Badge which has *GR*-value trend decreases, the *GR*-value trend of Winning Streak increases through each milestone. Moreover, we exploited that the number of users change their interesting from Badge to Winning Streak also increases. Therefore, we conducted the second analysis for discover the contribution of Winning Streak in entertainment improvement in Duolingo. By comparing the difference of *GR*-values between two user groups, the results indicated that the users, who has been attracted by Winning Streak, are more enjoyable rather than the normal users. The Winning Streak also works well in improving motivation of users and it is significant for advanced users rather than beginners. This analysis also be confronted with the real statistic data.

For the second research, we expressed the design of a cybersecurity awareness training game. In this study, we have used the activity theory-based model of serious games for designing our game. The story game play also used in this design. This game aims at giving players the real situation of cybersecurity problem. The situations is given in the main story and users will find the solution for those. By using Novelty, we have implemented our design as a game demo. This game demo has been played and evaluated by ten participants. The results of evaluations presented that the game was rated well on the aspect of understanding cybersecurity problems and solutions. It helps users avoid the rote learning and keep the knowledge in mind easily. Although the game can help users in understanding the cybersecurity concept, it was rated as not interesting. Therefore, we applied the game refinement idea to analyze and improve it. We also used gamification technique for creating learning structure for this game. After create a new structure. This game has *GR*-value falls into the appropriate range of sophisticated games.

Our current works showed the first approach in using game refinement measure for understanding and improving entertainment aspect of educational games. However, as we have stated, an assessment of educational game should be considered on two aspects which are entertainment and learning impact. Therefore, our future works will focus on figuring out the method which can give a common measurement for learning aspect. We hopefully expect that this thesis can create premise and inspiration for using a common measurement to understand and improve the effects of game elements in gamified systems.

Bibliography

- [1] Brandao Carvalho, M et al. “An activity theory-based model for serious games analysis and conceptual design”. In: (2015).
- [2] Brito, Jailson, Vieira, Vaninha, and Duran, Adolfo. “Towards a framework for gamification design on crowdsourcing systems: the GAME approach”. In: *Information Technology-New Generations (ITNG), 2015 12th International Conference on*. IEEE. 2015, pp. 445–450.
- [3] Camerer, Colin F. “Does the Basketball Market Believe in the Hot Hand?” In: *The American Economic Review* 79.5 (1989), pp. 1257–1261.
- [4] *Cybersecurity Awareness Game Demo*. 2017. URL: <https://www.dropbox.com/sh/xpz1fjmuqaljvxm/AAAGPRHfgZimUeWmQyx3QYiHa?dl=0>.
- [5] *Cybersecurity Awareness Game Survey*. 2017. URL: <https://goo.gl/forms/jZUDRFcx8TMdfsTJ3>.
- [6] Deterding, Sebastian et al. “From game design elements to gamefulness: defining gamification”. In: *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. ACM. 2011, pp. 9–15.
- [7] Diah, Norizan Mat et al. “Quantifying Engagement of Video Games: Pac-Man and DotA (Defense of the Ancients)”. In: *17th International Conference on Mathematical and Computational Methods in Science and Engineering (MACMESE15), Kuala Lumpur*. 2015.
- [8] *DOTA2-Hero kill*. 2017. URL: http://wiki.teamliquid.net/dota2/Gold#Hero_Kills.
- [9] Duolingo. *Learn a language for free*. (Accessed on 12/11/2017). URL: <https://www.duolingo.com>.
- [10] Fogg, Brian J. “A behavior model for persuasive design”. In: *Proceedings of the 4th international Conference on Persuasive Technology*. ACM. 2009, p. 40.
- [11] Fudenberg, Drew and Tirole, Jean. *Game theory*. Tech. rep. The MIT press, 1991.
- [12] Gee, James Paul. “What video games have to teach us about learning and literacy”. In: *Computers in Entertainment (CIE) 1.1* (2003), pp. 20–20.

- [13] Gené, Oriol Borrás, Núñez, Margarita Martínez, and Blanco, Ángel Fidalgo. “Gamification in MOOC: challenges, opportunities and proposals for advancing MOOC model”. In: *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*. ACM. 2014, pp. 215–220.
- [14] Gunter, Glenda A, Kenny, Robert F, and Vick, Erik H. “Taking educational games seriously: using the RETAIN model to design endogenous fantasy into standalone educational games”. In: *Educational Technology Research and Development* 56.5-6 (2008), pp. 511–537.
- [15] Hanus, Michael D and Fox, Jesse. “Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance”. In: *Computers & Education* 80 (2015), pp. 152–161.
- [16] Iida, Hiroyuki, Takeshita, Nobuo, and Yoshimura, Jin. “A metric for entertainment of boardgames: its implication for evolution of chess variants”. In: *Entertainment Computing*. Springer, 2003, pp. 65–72.
- [17] Iida, Hiroyuki et al. “An application of game-refinement theory to Mah Jong”. In: *International Conference on Entertainment Computing*. Springer. 2004, pp. 333–338.
- [18] Jeuring, Johan, Rooij, Rick van, and Pronost, Nicolas. “The 5/10 method: a method for designing educational games”. In: *International Conference on Games and Learning Alliance*. Springer. 2013, pp. 364–369.
- [19] Kananat, Suwanviwatana, Terrillon, Jean-Christophe, and Iida, Hiroyuki. “Gamification and Scrabble”. In: *Games and Learning Alliance*. Springer, 2016, pp. 405–414.
- [20] Kim, AJ. “Putting the fun in functional”. In: *Applying game mechanics to functional Software* (2008).
- [21] *Language Difficulty Ranking*. 2017. URL: <http://www.effectivelanguagelearning.com/language-guide/language-difficulty>.
- [22] McGonigal, Jane. *Reality is broken: Why games make us better and how they can change the world*. Penguin, 2011.
- [23] McMahon, Mark. “Using the DODDEL model to teach serious game design to novice designers”. In: *ASCILITE*. 2009, pp. 646–653.
- [24] MindSnacks. *MindSnacks*. (Accessed on 12/11/2017). URL: <https://www.mindsnacks.com>.
- [25] Moccozet, Laurent et al. “Gamification-based assessment of group work”. In: *Interactive Collaborative Learning (ICL), 2013 International Conference on*. IEEE. 2013, pp. 171–179.
- [26] Moon, Hyunkyung and Baek, Youngkyun. “Exploring variables affecting player’s intrinsic motivation in educational games”. In: *Proceedings of the 17th International Conference on Computers in Education*. 2009, pp. 718–722.

- [27] Nossal, Nathan and Iida, Hiroyuki. “Game refinement theory and its application to score limit games”. In: *Games Media Entertainment (GEM), 2014 IEEE*. IEEE. 2014, pp. 1–3.
- [28] O’Sullivan, F. J. *UDL Technology: Technology for Universal Design for Learning and Special Education*. 2016.
- [29] Pajak, Bozena. *Which countries study which languages, and what can we learn from it?* 2016. URL: <http://making.duolingo.com/which-countries-study-which-languages-and-what-can-we-learn-from-it>.
- [30] Sutiono, Arie Pratama, Purwarianti, Ayu, and Iida, Hiroyuki. “A mathematical model of game refinement”. In: *International Conference on Intelligent Technologies for Interactive Entertainment*. Springer. 2014, pp. 148–151.
- [31] Tan, Phit-Huan, Ling, Siew-Woei, and Ting, Choo-Yee. “Adaptive digital game-based learning framework”. In: *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts*. ACM. 2007, pp. 142–146.
- [32] Tang, Stephen, Hanneghan, Martin, and El Rhalibi, Abdennour. “Introduction to games-based learning”. In: *Games Based Learning Advancements for Multi-Sensory Human Computer Interfaces*. New York: IGI Global (2009).
- [33] *Visual Novel*. 2017. URL: https://en.wikipedia.org/wiki/Visual_novel.
- [34] *Visual Novel Maker*. 2017. URL: <http://www.visualnovelty.com>.
- [35] *WIKI-Winning Streak*. 2017. URL: [https://en.wikipedia.org/wiki/Winning_streak_\(sports\)](https://en.wikipedia.org/wiki/Winning_streak_(sports)).
- [36] Xiong, Shuo et al. “Quantifying engagement of various games”. In: *The 19th Game Programming Workshop 2014*. Information Processing Society of Japan. 2014.
- [37] Young, Michael F et al. “Our princess is in another castle: A review of trends in serious gaming for education”. In: *Review of educational research* 82.1 (2012), pp. 61–89.

Publications

Duy Huynh, and Hiroyuki Iida. “An Analysis of Winning Streak’s Effects in Language Course of Duolingo.” *Asia-Pacific Journal of Information Technology and Multimedia* 6.2 (2017).

Duy Huynh, Long Zuo, and Hiroyuki Iida. “Analyzing Gamification of Duolingo with Focus on Its Course Structure.” *Games and Learning Alliance*. Springer International Publishing, 2016. 268-277.

Duy Huynh, Phuc Luong, Hiroyuki Iida, and Razvan Beuran. “Design and Evaluation of a Cybersecurity Awareness Training Game.” In *International Conference on Entertainment Computing*, pp. 183-188. Springer, Cham, 2017.

Fatin Izzyan, **Duy Huynh**, Norshakirah Aziz and Hiroyuki Iida. “Comparative study: A Case Study in Analyzing Gamification between MindSnacks and Duolingo”. *ICIDB2018*.

Duy Huynh, Long Zuo and Hiroyuki Iida. “An Assessment of Game Elements in Language-Learning Platform Duolingo”. *ICCOINS2018* (submitted)