HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Language Education
English Language Teaching Program

OCTALYSIS İLE DİL ÖĞRENME ORTAMLARININ OPTİMİZE EDİLMESİNDE OYUNLAŞTIRMA VE YÖNLENDİRİLMİŞ MOTİVASYON AKIMLARI

Mehmet ABİ

Ph.D. Dissertation

Ankara, (2021)

With leadership, research, innovation, high quality education and change,

To the leading edge... Toward being the best...

HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Language Education
English Language Teaching Program

GAMIFICATION AND DIRECTED MOTIVATIONAL CURRENTS IN OPTIMIZING LANGUAGE LEARNING ENVIRONMENTS THROUGH OCTALYSIS

OCTALYSİS İLE DİL ÖĞRENME ORTAMLARININ OPTİMİZE EDİLMESİNDE OYUNLAŞTIRMA VE YÖNLENDİRİLMİŞ MOTİVASYON AKIMLARI

Mehmet ABİ

Ph.D. Dissertation

Ankara, (2021)

Acceptance and Approval

To the Graduate School of Educational Sciences,

This dissertation, prepared by **MEHMET ABİ** and entitled "Gamification and Directed Motivational Currents in Optimizing Language Learning Environments through Octalysis" has been approved as a thesis for the Degree of **Ph.D.** in the **Program of English Language teaching** in the **Department of Foreign Language Education** by the members of the Examining Committee.

Chair	Prof. Dr. Nuray Alagözlü
Member (Supervisor)	Dr. Öğr. Üyesi Hatice Ergül
Member	Prof. Dr. Hacer Hande Uysal Gürdal
Member	Prof. Dr. Eda üstünel
Member	Prof. Dr. Şevki Kömür

This is to certify that this thesis/dissertation has been approved by the aforementioned examining committee members on …/…/…. in accordance with the relevant articles of the Rules and Regulations of Hacettepe University Graduate School of Educational Sciences, and was accepted as a **Ph.D. Dissertation** in the **Program of English Language Teaching** by the Board of Directors of the Graduate School of Educational Sciences from ...../...../.....

Prof. Dr. Selahattin GELBAL
Director of Graduate School of Educational Sciences

Abstract

For a long time, games have been used to increase individuals' engagement with the tasks in non-game contexts. However, it is seen that integrating a whole game into a context or creating a totally new game for a specific context can cause several problems in practice. Thus, researchers, especially the ones who are interested in digital games, have proposed game elements, which enable users to go on playing and make a game fun, to be used in non-game contexts, and have developed the concept of gamification. And scientists from various fields have conducted studies on gamification since 2010, and developed many models. Yet, when investigated, it is seen that their studies lack statistical evidence to validate their models. In this study, theoretical background of Octalysis which is one of the comprehensive gamification models was investigated in terms of Directed Motivational Currents Theory and whether gamification could be used for this purpose was tested. Hypothesis were formed according to the factors which are told to affect motivation and their relationships with long-term motivation and they were tested through Structural Equation Modeling. In the second phase of the study, an Octalysis application was applied in face-to-face foreign language classrooms and the model was tested empirically. Qualitative and quantitative data which were obtained during the intervention were compared and a justification about usability of the model was made. As a result, it was found out that gamification applications could be used to improve students' engagement with tasks and/or classes to improve students' motivation.

Keywords: gamification, octalysis, flow theory, self-determination theory, directed motivational currents.

Öz

Oyun dışı ortamlarda kişilerin daha çok motive olmaları ve aktivitelere katılımlarının arttırılmaları amacıyla oyunlar uzun zamandır kullanılmaktadır. Ancak bir oyunun tamamen bir ortama entegre edilmesi ya da özel bir ortam için yeniden bir oyun oluşturulmasının uygulama bakımında sıkıntılar oluşturduğu görülmüştür. Buradan yola çıkan ve özellikle sayısal ortamlar için oyunlarla ilgilenen araştırmacılar, oyunun bir bütün olarak kullanılmasından ziyade oyunu eğlenceli yapan ve kişilerin oynamaya devam etmelerini sağlayan oyun unsurlarının oyun dışı ortamlarda kullanılması fikrini ortaya atmışlar ve oyunlaştırma kavramını geliştirmişlerdir. Çeşitli alanlardan bilim adamları özellikle 2010 yılından itibaren oyunlaştırma üzerine çok sayıda çalışma yapmışlar ve oyunlaştırmanın daha verimli uygulanabilmesi için birçok model geliştirmişlerdir. Ancak çalışmalarda modellerin geçerliliklerini doğrulayacak istatistiksel bulguların eksik olduğu görülmektedir. Bu çalışma kapsamında oyunlaştırma alanında geliştirilen modellerden birisi olan Octalysis' in kuramsal altyapısı incelenmiştir. Ayrıca, Octalysis modelinin dayandığı teorik kavramlar, yabancı dil eğitiminde son yıllarda öne çıkan Yönlendirilmiş Motivasyon Akınları Teorisi bağlanımda ele alınmış ve oyunlaştırmanın bu amaçla kullanılıp kullanılamayacağı test edilmiştir. Modelde motivasyonu etkilediği belirtilen faktörlerin uzun vadeli motivasyonla ilişkileri ile ilgili hipotezler oluşturulmuş ve bu hipotezler yapısal eşitlik modellemesi yoluyla test edilmiştir. Elde edilen veriler doğrultusunda araştırmanın ikinci aşamasında Octalysis modeline göre yüz yüze eğitim yapılan bir yabancı dil sınıfında oyunlaştırma uygulaması yapılmış ve model deneysel olarak test edilmiştir. Ayrıca, uygulama sırasında nitel araştırma yöntemleri kullanarak elde edilen veriler ile nicel veriler karşılaştırılmış ve modelin kullanılabilirliği ile ilgili bir yargıya varılmaya çalışılmıştır. Araştırmadan çıkan sonuçlara göre, yüz yüze yabancı dil eğitiminin uygulandığı sınıflarda öğrencilerin derslere ve/veya aktivitelere katılımlarının ve motivasyonlarının arttırılmasında oyunlaştırma uygulamalarının kullanılabileceği saptanmıştır.

Anahtar sözcükler: oyunlaştırma, octalysis, akış teorisi, öz-belirleme teorisi, yöneltilmiş motivasyon akıntıları teorisi.

In the memory of my beloved father Salih Alper Abi.

Rest in peace ...

Acknowledgements

This study would not have been possible without the help of many people. First of all, I would like to take this opportunity to express my immeasurable gratitude to my supervisor Dr. İsmail Hakkı Erten who guided and supported me a lot at the beginning of this journey. I have been privileged to know him not only as my supervisor and teacher but a brother and a friend. His untimely death tore our hearts out. Even though he is gone, he is still with us. We will all remember him as a gentleman, a guide, and a humble man. I also would like to remember Dr. Hüseyin Öz whose sudden loss brought nothing but grief. He always did not refrain to support me starting from the first day I met him and encouraged me.

After losing Dr. Erten, I was despaired. But my second advisor Dr. Hatice Ergül whose support it would not be possible to come to this point. I wish to express my gratitude to her. I am also thankful to Dr. Eda Üstünel and Dr. Şevki Kömür who introduced me into ELT. Without their guidance and inspirations, I would not have develop a perspective in this field. It is a genuine pleasure to know them and work with them.

Special thanks to friends Yeşim Türkekul who was very patient and understanding during all my studies; Ali Aktaş whose generous assistance and tolerance was meaningful and priceless. I would also thank to my friend Ahmet Taşpınar for his endless support, friendship and great hospitability.

My genuine thanks go to my fellow colleagues in School of Foreign Languages in Muğla Sıtkı Koçman University and Dr. Mete Ersoy High School/Milas for their contributions and support.

Last but not the least I would like to thank my wife Elif Abi who always encouraged me and put up with my crazy schedule. Without her, this study would not be possible. I am greatly indebted to my daughters, Azra Abi and Alya Abi, who sacrificed a lot. But it was their unconditional love and tolerance which made this process worthwhile.

Table of Contents

List of Tables

List of Figures

Symbols and Abbreviations

PBLs: Points, badges and leaderboards

SLA: Second language acquisition

ELT: English Language Teaching

L1: First/Mother language

L2: Second/Foreign language

EFL: English as a foreign language

CD: Core drive

DMC: Directed motivational currents

EFA: Exploratory factor analysis

CFA: Confirmatory factor analysis

SEM: Structural equation modelling

ICT: Internet and communication technologies

MOOCs: Massive open online courses

NMC: New media consortium

XP: Experience point

CEFR: Common European framework of reference for languages

SDT: Self-determination theory

MCAR: Missing completely at random

MSKU: Muğla Sıtkı Koçman University

SFL: School of foreign languages

CEI-II: The Curiosity and exploration inventory II

BPNSG-General: Basic psychological need satisfaction and frustration scale – general measure

AGQ-R: Achievement goal questionnaire revised

DMCDQ: Directed motivational currents disposition questionnaire

POQ: Psychological ownership questionnaire

STPS: Susceptibility to persuasion scale

SPSS: Software program for statistical analyses

AMOS: analysis of moment structures

MSA: Measure of sampling adequacy

ML: Maximum likelihood

GoF: Goodness-of-fitness

GOF: Goodeness-of-fit

RMSEA: Root mean square error of approximation

RMR: Root mean square residual

SRMR: Squared root mean square residual

NFI: Normed fit index

CFI: Comparative fit index

RFI: Relative fit index

PNFI: Parsimony fit index

AGFI: Adjusted goodness-of-fit index

AVE: Average variance extracted

CR: Composite reliability

HTMT: Heterotrait monotrait

MI: Modification index

KMO: Kaiser-Meyer-Olkin

CMIN/DF: Relative chi-square

CMIN: Chi-square

SE: Standard error

EM: Estimated means

Chapter 1
Introduction

This section consists of the statement of the problem, the aim and the methodology of the study, research questions, premises, limitations and definitions.

Statement of the Problem

Developments in computer and internet technologies have caused radical changes in our lives and many things are attuned to these changes. It can be said that games are not exceptions in this sense. Through these changes, it has become possible to play many games in virtual environments whereas they used to be only played outside or inside physical environments before (Şahin & Samur, 2017). Although there are still some classical games which have resisted the destructive nature of time and preserved their initial forms, people mostly prefer playing games online on their mobile devices or personal computers. This, as Yılmaz and Çağıltay (2004) point out has led the emergence of a new sector – game sector.

To increase the popularity of the virtual games and enable people to play more, the game sector has begun to investigate games more closely and try to explore what drives people to play games; what motivational factors play roles in engaging them and how. Soon, researchers have begun to realize that games share some common characteristics which determine their effectiveness. In terms of why games are so attractive, researchers such as Kirriemuir and McFarlane (2004) have found out that fantasy, challenge and curiosity are among the main factors underlying games' success. As studies go on, researchers have started to group these characteristics and named them as elements, mechanics and techniques. it has been discovered that game elements and mechanics are highly efficient in motivating players for desired behaviors (Bunchball, 2010; Werbach & Hunter, 2012).

Consequently, game designers have begun to integrate game elements and mechanics into the game design processes. It has been found out that games designed to include specific game elements and mechanics are being very in terms of motivating people and getting them engaged. For example, in her TED talk, McGonigal (2010) mentioned the game World of Warcraft which has been designed in accordance with the game elements and mechanics. The number of active users

of World of Warcraft is 5.93 million, showing how effective game elements and mechanics can be when they are used systematically, and this can exemplify the power of games.

Having realized the success of video games in virtual environments in motivating users and getting them engaged, researchers from various fields have wanted to see whether the game elements that make these games successful in virtual game environments can be applied to other contexts. As a result, they have developed the concept of “gamification”, which means using game elements in non-game contexts (Deterding, Sicart, Nacke, O’Hara, & Dixon, 2011).

Gamification, which was first used in 2008, has begun to be used extensively only after 2010 (Deterding et al., 2011). Soon, it has attracted the interest of many researchers from various fields and become the topic of many studies. As in many fields such as economy, finance, marketing, medicine and so on, researchers in the field of education have not remained indifferent to gamification and they have started to investigate the potential use of game elements in educational contexts which are accepted as non-game contexts.

When the studies on gamification are investigated, it is possible to see that the studies that were conducted in the first phase of the gamification history focused on what gamification was and why it should be applied while the studies in the second phase dealt with how and when gamification should be used or how or when it should not be used (Nacke & Deterding, 2017). In other words, in initial studies what game elements were and how these could be employed for gamification and for what purposes it should be used were the main goals. Like many other researchers, Brathwaiter and Schreiber (2008), Bunchball (2010), and Giannetto, Chao and Fontana (2013) deal with these issues in their studies. According to these researchers, among the most frequently used game elements used in gamification applications are points, badges and leaderboards (PBLs), feedback and rewards.

As it can be guessed, many studies have been conducted on the use of game elements in gamification applications. In addition, a great number of these studies have found out that it is possible to obtain positive results with gamification (Werbach & Hunter, 2012). However, as the scope of the studies on gamification expands, it is seen that some researchers such as Reeves and Read (2009) and

Hanus and Fox (2015) have adopted rather critical point of view. Researchers from these group advocate that using game elements as extrinsic motivators may not always have positive impact or it may not be possible to reach clear conclusions in terms of the positive effects of gamification applications.

Generally speaking, when studies carried out so far are investigated, it is seen that researchers have similar standpoints in terms of positive effects of gamification on motivation and engagement. For example, Jung, Schneider, and Valacich (2010), and Mekler, Brühlmann, Opwis and Tuch, (2013) state that gamification is highly effective in reaching the desired goals in non-game contexts.

It is possible to see similar findings in studies carried out in educational contexts. Kapp (2012) points out that games have been in use in educational contexts for many years and researchers including Michael and Chen (2006), Gee (2014), Di Bitonto, Corriero, Pesare, Rossano, and Roselli (2014), and Nacke and Deterding (2017) remark that using games in educational environments can provide positive outcomes. However, integrating games as a whole into learning and teaching processes or creating totally new games specific to educational purposes can bear some problems (Simoes, Redondo, & Vilas, 2013).

Thus, researchers have begun to think that gamification can be a solution for the problems resulting from using games themselves. Thus, they have started to concentrate on using game elements and gamification for educational purposes. Researchers like Dicheva, Dichev, Agre and Angelova (2015) and Corbett (2010) carried out studies on how to apply gamification. In this vein, although they figure out that gamification in education has been mostly applied in virtual environments such as mobile learning or e-learning courses, it is seen that there are several studies on gamification of physical classrooms. Hew, Huang, Chu, and Chiu (2016) and Giannetto et al. (2013) and many other researchers argue that gamification in education is beneficial and can positively affect students' motivation and engagement levels.

When the application of gamification in English as a foreign language (EFL) classrooms is investigated, surprisingly it is seen that there are relatively few studies using gamification. According to the findings from Baber's (2015) literature review in 2015, there are only nine studies in EFL which employ gamification whereas this

number is over thousands in other educational fields. There are already many applications which can be accessed through mobile devices and which aim at providing language courses such as Memrise, Duolingo, Quizfun and so on (Berg, 2013). On the other hand, there are a few studies carried out in physical EFL classrooms. For example, Baber (2015) proposes two types of scenarios to be used in language classrooms which are designed by using gamification approach.

As to the gamification studies in Turkish contexts, it is seen that a similar tendency has been adopted by Turkish researchers. Especially in educational contexts, many researchers such as Birant (2014), Bozkurt and Genç-Kumtepe (2014), Çağlar and Kocadere (2015), Erdoğdu and Karataş (2016) have carried out studies on gamification. Turan, Avinc, Kara, and Göktaş (2016), and Güler and Güler (2015) and other researchers have found out that gamification has positive effects on motivation and engagement.

In this respect, some researchers such as Aleven, Myers, Easterday and Ogan (2010), Rapp (2017), and Aparicio, Vela, Sanchez, and Montes (2012) have proposed some models and frameworks to apply gamification in non-game contexts. Huang and Soman (2013) state that by using these models it can be possible for teachers to re-design their classes and make them more productive. A framework at the center of which is the Self-Determination Theory (Deci & Ryan, 2000) is developed by Aparicio et al. (2012). According to them, when gamification is designed following a framework, gamification can be applied more effectively. Therefore, first the goals should be determined and adopted and then, the game mechanics to be used should be determined. Finally, the framework should be evaluated.

There are many frameworks and models proposed by many researchers. However, one of the prominent models in recent years is proposed by Chou (2015). According to this comprehensive model which is called as “Octalysis”, there are eight Core Drives (CDs) underlying people’s behaviors. That is, people act due to at least one of these eight CDs. Thereby, gamification applications should be carried out by taking these drives into consideration. If not, employing game elements aiming at extrinsic motivation such as PBLs cannot be successful. One distinctive feature of Chou’s model is that it has a different perspective beyond points, badges and leaderboards. Instead of utilizing individual game elements, he suggests using

game techniques which consist of different game mechanics. Another strength of the Chou's model is its ability to combine different theories into one model. His model mainly comprises Self-Determination Theory (Deci & Ryan, 2000), Flow Theory (Csikszentmihalyi, 1996), Approach and Avoidance Motivation (Elliot, 1999), Goal Setting Theory (Locke & Latham, 1990). All these theories have attracted many researchers from different fields. Although Chou (2015) puts the Flow Theory at the center of his model, in scope of this study, Directed Motivational Currents (DMC) Theory by Dörnyei, Henry, and Muir (2016) is used. Actually, DMC Theory is not different from Flow Theory. In contrast, it compensates the weak sides of Flow Theory.

Another powerful side of Octalysis is that it can allow teachers take a kind of picture of their classrooms and diagnose the problems causing demotivation or decreasing students' engagement levels. In this respect, it is possible to overcome some of the criticisms made by researchers like Landers and Callan (2011), and Landers and Armstrong (2017) who question the use of game elements as gamification, which might result in superficial applications. Because gamification needs to be applied in accordance with the requirements defined by the experts. Therefore, it can be said that Octalysis can be used as a diagnostic tool with which specific game techniques which are used to solve problems can be chosen easily.

Thus far, it can be understood that gamification can positively affect motivation and enable obtaining desired behavioral changes. However, it may not be correct to say that gamification studies are free from some drawbacks. One of the biggest criticisms directed is that proposed models and frameworks lack studies testing the validity of them statistically. Seaborn and Fels (2015) stress that in many gamification studies, theoretical background is not reinforced enough and interpreted sufficiently and thus, there are gaps between theory and practice. Besides, more experimental studies are needed to overcome this problem. That is, as Nacke and Deterding (2017) indicate, since application areas are various and too many factors play role, it is evident that gamification should be founded on sound and valid theoretical foundations. Also, gamification applications should be designed by taking the characteristics of target activities into consideration in a specific context.

When the aforementioned data are taken into consideration, it is thought that in EFL classrooms, gamification applications can be used to increase students' motivations and engagement levels. It is clear that it would be beneficial to use game techniques or game mechanics according to the diagnostic data which will be obtained by using Octalysis in EFL classrooms. In addition, for sound theoretical foundations, there is a need for experimental studies.

Aim and Significance of the Study

It is indicated by the researchers, gamification approach which means using game elements in non-game contexts (Deterding et al., 2011) may have positive impact on students' motivation and engagement in educational contexts (Di Bitonto, 2014) such as EFL classrooms and online learning environments (Baber, 2015). In addition, various models have been developed by some researchers (Aparicio et al., 2012; Chou, 2015; Nicholson, 2012; Rapp, 2017; and Sakamoto, Nakajima, & Alexandrova, 2012) to enable gamification to be applied in more structural ways. Among these models "Octalysis" which is developed by Chou (2015) comes to the fore as a prevailing tool because it is possible to integrate major motivation theories into Octalysis. Moreover, it can be used as a kind of diagnostic tool in classrooms and can be used to identify the weak sides of classrooms which require intervention. Another benefit is that there is no need to ignore the individual differences in classrooms since Octalysis is a human centered model and enables making plans according to the different needs of different students. Besides, by using Octalysis, it is also possible to make designs which take initial attractor states into consideration, which means that the dynamic perspective of classrooms can be integrated into the system. Although these make Octalysis an influential tool, perhaps its most powerful side is its different approach to game elements because Chou (2015) suggests that gamification research should go beyond mere use of PBLs. Instead, the drives underlying games which make games so attractive and motivating should be focused on in design processes. In this respect, Chou suggests several game techniques which include game elements and game mechanics.

However, it is possible to talk about similar criticisms toward Octalysis as to other game models. As Seaborn and Fels (2015) point out, most of the gamification studies so far have failed to offer sound theoretical backgrounds. In addition, there

are gaps between theory and practice. Therefore, more experimental research is needed.

The primary aim of this study is to identify whether the eight CDs in Octalysis can be used in EFL classrooms or not. In addition, whether Directed Motivational Currents Theory (Dörnyei, et al., 2016) can be integrated into the system instead of the Flow Theory (Csikszentmihalyi, 1990) since the Flow Theory, which is used at the center of the model, is found to be problematic by some researchers such as Dörnyei et al., (2016). Thus, in the current study, it is aimed to develop a scale by utilizing several valid and approved scales which are in accordance with the eight CDs of Octalysis and to test this scale statistically. In this way, it may be possible to determine the validity of Octalysis as a gamification model in EFL classrooms. The current study also aims to design a gamification application according to the results of the scale and to test this design empirically. As a result, it is expected to reveal sound findings about gamification in EFL classrooms and provide a useful tool that can be used by EFL teachers.

Research Questions

In the current study, it was aimed at discovering whether it was possible to utilize Octalysis framework to improve long-term motivation in language learning environments which is theorized as Directed Motivational Currents by Dörnyei et al. (2016). In accordance with this aim the main research question of the study was:

- Can DMC integrated Octalysis as a gamification model be used in EFL classrooms to enable sustainable long-term motivation?

In the first phase of the study the goal was to find out whether the adapted versions of the preselected tools which were thought to measure the eight dimensions of the Octalysis framework could be used as suggested by their authors. The research question for this stage was:

- Can preselected scales be used as suggested by their authors to measure the eight core drives of Octalysis?

Following the EFA procedures, it was aimed to investigate how well the measured variables could measure the constructs underlying Octalysis. That is, the goal was

to find whether there was a consistency between the data and the measured constructs. The research question for this stage was:

- Is it possible to measure what is intended to measure with the adapted version of the survey?

In addition, in the scope of this study the following hypothesis were formed and tested:

English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with:

1. their feeling of accomplishment, skills development and overcoming challenges
2. their involvement in creative processes
3. their feeling of the ownership regarding their own learning processes
4. their feeling of being related to a social group in their learning environments
5. their desire to obtain scarce things and their perception of closing deadline
6. their being curious about the things in their learning environments which get their attention and their encounter with unpredictable things
7. their endeavors not to lose something and to avoid failure

The secondary objective of the study was to find out whether a gamification intervention based on Octalysis theory could contribute to the improvement of DMC in English as a foreign language classes, which was the objective of the second phase of the study.

For this purpose, two classrooms were randomly selected by the researcher, one of which was called as experimental group and the other as control group. First, the developed survey was carried out in written format with the participation of volunteer students. The data obtained at this stage was coded and independent sample t-

tests were carried out to examine whether there are any differences between the two groups. The research question formed at the stage was:

- Are there any differences between the experimental and control groups before the implementation of gamification intervention?

Then, with semi-structured audio-recorded interviews were conducted. With these interviews it was aimed to diagnose the motivation levels of participants in language learning environments and then draw the Octalysis frameworks that illustrated those problems on the octagon. By examining the Octalysis frameworks it, then, could be possible to detect any significant differences between groups. The research questions related to this stage were:

- Can semi-structured pre-intervention interviews provide data to draw Octalysis frameworks of experimental and control groups?
- Is it possible to identify motivational drives by using pre-intervention Octalysis frameworks in experimental and control groups?

Finally, with the post-intervention, it was aimed to find out the effectiveness of the DMC integrated Octalysis gamification interventions in English as a foreign language classrooms. The research questions for this stage were:

- How is the DMC integrated Octalysis gamification intervention perceived by the participants in the experimental group?
- Are there any differences between the experimental and control groups after the intervention?

Assumptions

Before starting the study, there were some assumptions to be considered. First of all, it was not possible to carry out the study with the whole target population. Thus, it was assumed that the samples were the representatives of the population. Another thing was that although the selected scales and questionnaires had already been tested and validated by their authors, the assumption was that all the tools utilized within the scope of the current study were in line with the purposes of the study. In addition, they were valid and reliable tools since they were developed by experts.

Within the scope of the study, all the participants who had taken part in the various phases of the study were informed about their rights and the aims and the procedures to be carried out, and asked to give their consents in written format and in some cases both in written and oral format; thus, all participants were assumed to take part in the study voluntarily and willingly. In the first phase of the study consisted of three studies, the participants were asked to indicate their levels of agreement to the Likert-type statements. It was assumed that they stated their opinions truthfully and honestly.

In the second part of the study, they were asked to fill out the survey structured in accordance with the findings obtained in the first phase. Then before the gamification implementation, the participants (students and teachers) were invited to take part in semi-structured audio-recorded interviews and asked to evaluate their language learning experiences in accordance with Octalysis theory. Based on their answers, the gamification implementation was structured. After the implementation, they were again invited to take part in semi-structured audio-recorded interviews to evaluate the effectiveness of the gamification intervention. In all these data collection steps, it was presumed that participants responded honestly and decently.

Finally, the findings of the study were believed to provide answer to the primary question of whether it could be possible to utilize gamification theory of Octalysis to increase long-term motivation (DMCs) in language learning classrooms.

Limitations

As in many studies, this study was not free from some limitations. One of the limitations was that the data collected throughout the study was limited only to the participants' views.

In the first phase of the study, the data was collected via selected scales which were aimed at identifying underlying structures of Octalysis framework. The items in the scales were adapted for the research purposes and were expected to be the representatives of the latent structures of Octalysis. In this sense, the high number of the items used in the scales could be seen as a limitation. However, it was necessary to include so many items in to the Octalysis survey because Octalysis framework consists of eight core drives which are believed to affect

students' levels of motivation in language learning environments in various ways. Thus, to cover all the dimensions of Octalysis, the number of the items included in the survey was over 50. When the fact that the multivariate statistical procedures were carried out in the scope of the first phase of the study is taken into consideration, the close relationship between the number of items and required sample sizes could be seen as a limitation. To overcome this problem, EFA procedures were carried out three times with the participation of different populations each time. Then, CFA and SEM procedures were conducted.

In the first part of the study, all the participants were language learners at tertiary level and attending compulsory English language preparatory classes at B2 level in six universities across Turkey. So, they were supposed to have similar language proficiency levels, and the data collected in Studies 1, 2 and 3 were limited to this population which was a portion of language learner population in Turkey.

In the second phase of the study, the number of participants was limited to the volunteered students and teachers (10 students and 2 teachers). In addition, they were not aware of the concept of gamification and Octalysis framework. Thus, to prevent any misconceptions, prior to the semi-structured interviews, briefings were given both to the students and teachers about the Octalysis framework. Besides, all the interviews in this phase were carried out in native language of the participants which is Turkish to enable them to express their thoughts more clearly and thoroughly. Their audio-recorded responses were then translated into English by the researcher, and presented in the study in English.

Finally, within the context of this study, only Level 1 Octalysis principles were utilized. Indeed, in the original designs of Octalysis are applied at four levels. In each level, the complexity of the design increases and different layers are added to the designs. For example, at the first level, the different phases of learning motivation were taken into consideration whereas at subsequent levels, learner types and their attractor states were included, which are among the interest of dynamic systems theories of motivation. In this sense, level 1 application of Octalysis could be another limitation.

Definitions

Directed motivational current: An intense motivational drive – or surge – which can stimulate and support long-term behavior.

Eudemonic well-being: A deep and often enduring sense of personal contentment.

Extrinsic motivation: motivation which occurs when people do something because of the outcomes.

Flow experiences: The experiences that occur when one is totally involved in a task at hand.

Gamification: Using game elements in non-game contexts

Game: A system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome

Game mechanic: An element of a game that is made up of a set of rules and feedback loops used to incentivize the player.

Game element: Elements that make up games

Game dynamics: The interactions between game mechanics and players

Serious games: Games designed for any purpose rather than entertainment

Game-based-learning: A type of game play that has defined learning outcomes

Gamification of education: The use of game elements in a learning environment, usually with the support of ICT

Intrinsic motivation: Motivation resulted from inherent interesting or enjoyable nature of the task at hand.

Positive psychology: The scientific study of what enables individuals and communities to thrive.

Self-concordant goals: Goals determined by the individual himself/herself and represent individual's interests, passions, values, and beliefs.

Chapter 2
Literature Review

Gamification

It will not be wrong to say that gamification is a relatively new research area which is getting the interest of more and more researchers from various fields. The developments in web technologies and digital business models and online and location-based gaming result in the emergence of gamification research (Nacke & Deterding, 2017). Throughout this process, different game types have been investigated. And researchers seek better ways of applying different games for motivational purposes in different contexts. Rapp (2017) summarizes this process and indicates that as a result of the studies on serious games, causal games and pervasive games, researchers can apply game elements in various fields in daily life which are non-game context.

Games and gamification. It is evident that researchers have tried to make use of games in various non-game environments. To better understand what has given way to this, it is important to differentiate between the two major terms used in relation with this topic: play and game. And it is crucial to understand what is a play and what is a game (Deterding, 2011) and the effects of games because when the issue of what makes a game game is clearly defined, it may become possible to employ these in non-game contexts.

To understand how effective games can be in terms of increasing motivation, McGonigal's (2010) example which she gave in a TED talk can be useful. It was stated that the well-known game called World of Warcraft which was launched in 2010 has been played by 5.93 million people. This huge number can prove how effective game-designs can be in terms of motivation. In this respect, researchers like Schell (2008) try to identify some characteristics of games. In that, games provide feedback, create a sense of progress and possibility of success. In addition, games are engaging and they trigger curiosity and allow players feel free to make their own choices to overcome challenges. In another study, Juul (2003) identifies six main features common to all games which are: rules, variable, quantifiable outcomes, value-laden outcomes, player effort, player investment, and negotiable consequences.

So how can the term game be defined? Salen and Zimmerman (2004) define a game as “a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome” (p. 80). When there are clear goals and outcomes, the activity can be called as a game. Kuo and Chuang (2016) state that in games there are sets of rules, activities towards achieving a goal; however, when the definitions for play are examined, a different approach is seen. According to Garvey (1990), the main goal of play is fun and enjoyment and participants take part in play voluntarily and mostly they are intrinsically motivated during plays. When these definitions are compared, it is seen that what differs games from plays is their goal and measurable outcomes. In this respect, as Kuo and Chuang (2016) argue, games can be seen as exterior expressions of plays which means they are sources of joyful and intrinsic motivation.

In the process of gamified interventions, extrinsic motivators can be utilized to increase engagement in non-game contexts in which people are generally not intrinsically motivated. From a similar perspective, Kirriemuir and McFarlane (2004) argue that games are attractive because they have the power of merging fantasy, challenge and curiosity. In addition, they provide appropriate contexts in which individuals can enter the ‘flow’. Thus, as De-Marcos, Dominguez, Saenz-de-Navarrete and Pages (2014) put forward, game-thinking and playful designs are used as motivational tools to increase engagement in non-game contexts.

Defining gamification. Having seen that some of the game components can be utilized in non-game contexts, researchers have developed gamified interventions through which it is aimed to integrate more playful components in non-game contexts. This phenomenon is called as gamification by Deterding et al. (2011). Although, the term ‘gamification’ is first used in 2008, people have started to use it only after 2010 (Deterding et al., 2011).

Despite its short history, gamification which is at first a new research topic has grown into a multi-disciplinary field (Nacke & Deterding, 2017). Although the field has attracted many researchers’ interests, the number of game elements and theories that have been investigated so far is quite limited. Thus, researchers conclude that studies that have been designed to uncover psychological mediators and behavioral outcomes are needed. In terms of design elements in the first phase of gamification research, today it is possible to see researchers who are trying to

design new interventions by taking different game elements into consideration and focusing on motivational foundations. But as Nacke and Deterding (2017) point out there is still a need to combine research and practice to evaluate the usefulness of designs.

When the historical development of gamification is examined, two phases emerge (Nacke & Deterding, 2017). Accordingly, in the first wave, researchers tried to find definitions and developed frameworks and taxonomies for gamification and game design elements. In addition, they mostly dealt with describing systems, designs and architectures. Effect and user studies of gamified systems were also examined. In other words, researchers in this first wave tried to find answers for questions “what?” and “why?”. But recent studies are looking for answers for “how?”, “when?” and “how and when not?”.

So far, several researchers (Deterding et al., 2011; Houtari & Hamari, 2012; Seaborn & Fels, 2015; Werbach, 2014) have tried to define what gamification is. However, they have failed to provide an unanimously accepted definition for gamification. Zichermann and Cunningham (2011, p. xiv) define gamification as “the process of game thinking and game mechanics to engage users and solve problems”.

One of the mostly used definition of gamification is made by Deterding et al. (2011, p. 2). They define gamification as the “use of game mechanics in non-game contexts”. To better understand what is meant by this definition, maybe it would be useful to look at what is meant by non-game contexts. According to Deterding et al. (2011) and Groth (2012), entertainment is the primary aim of games and when this is taken into consideration, it can be said that any context in which the main purpose is not to entertain can be accepted as non-game context.

These definitions focus on the use of certain game mechanics for gamification and this allows researchers to be able to apply gamification in various contexts. On the other hand, limiting gamification to the integration of game mechanics may cause several problems. As Nacke and Deterding (2017) assert “at the heart of gamification design process is the development of gameful systems, which are complex combinations and interactions between elements” (p.453), and studies which just focus on the role of individual elements may not facilitate the

understanding of such complex systems. Interaction of game design elements and the dynamics that emerge during the application should be investigated.

In search of more practical definitions, researchers focus on the motivational factors that can be found in games and application of gamification accordingly. For example, Houtari and Hamari (2012) define gamification as “a process of enhancing services with (motivational) affordances in order to invoke gameful experiences and further behavioral outcomes”. They emphasize that by applying gamification, it is possible to invoke the same experiences as games do. In a more practical perspective, Koivisto and Hamari (2014) state that gamification is “the phenomenon of creating gameful experiences” (p.174). Werbach (2014) similarly defines it as “the process of making activities more game-like” (p.6).

In spite of these efforts, the multi-faceted and dynamic nature of gamification, which can be applied in various contexts, causes further problems in terms of reaching appropriate definition. This view is also shared by Seaborn and Fels (2015) who state that there is not any agreed upon definition of gamification. However, in most sources it is generally accepted that gamification includes use of game elements and mechanics in non-game contexts. The authors give a standard definition of gamification as “the intentional use of game elements for a gameful experience of non-game tasks and contexts. Game elements are patterns, objects, principles, models, and methods directly inspired by games” (p. 17). With a similar approach, Hamari, Koivisto and Sarsa (2014) conceptualize gamification in three parts: implemented motivational affordances, the resulting psychological outcomes and the further behavioral outcomes (p. 3026).

When the application of gamification is examined, it is seen that the issue is not so simple. Since different contexts have different characteristics, it may not be possible to find a one-size-fits-all application. The individual differences in each context and their effects on the design processes in different contexts should still be investigated. As Nacke and Deterding (2017) state “we need validated theories how design elements function and interact with individual dispositions, situational circumstances, and the characteristics of particular target activities” (p. 453).

Game elements, mechanics and designs. In addition to understanding what is game and what is play, some common terms frequently used in gamification

such as game elements, game mechanics, game dynamics, game designs and so on should be investigated for better understanding.

In terms of the relationship between game elements, game design and game mechanics, Giannetto, et al., (2013) state that only specific elements of game systems can form gamification. When these elements are merged together with idea of 'design', then, they become to be referred as game mechanics. In other words, game elements are accepted as the sub-category of game mechanics. Gåsland (2011) defines the term game mechanic as "an element of a game that is made up of a set of rules and feedback loops used to incentivize the player" (p. xiii). It is understood that game mechanics have motivating roles. Many things such as items, points, levels and bonuses can be counted as game mechanics. As it is seen, the terms game mechanics and game elements occasionally are used interchangeably. To clarify this point Brathwaiter and Schreiber (2008) use the term game design atoms to indicate basic game elements. They point out that game mechanics which are parts of game design are above the game elements and can change the game state.

In terms of the relationship with game mechanics and game dynamics, Bunchball (2010) states that game mechanics are related to the rules and features of a game such as fun, rewarding or any other desired emotions. The wishes and motivations that result in these emotions are called game dynamics. Zichermann and Cunningham (2011) point out that game mechanics are related to the functioning of game elements whereas game dynamics are the interactions players have with the mechanics. Seixas, Gomes, and de Melo Filho (2016) state that although the terms game mechanics and game dynamics are used as synonyms, they are actually different terms. Some of the most common game mechanics are points, levels, challenges trophies, badges/medals and accomplishments, virtual goods and classification tables, ranking and score table (Bunchball, 2010). According to Bunchball (2010), game mechanics motivate because of their dynamics. They have some needs and desires. Werbach and Hunter (2012) state that game mechanics such as challenge, rewards, feedback can be used to create engagement and involve essential processes. On the other hand, game dynamics such as constraints, emotions, progression and relationship are not directly included in the process. Figure 1 shows the relationship between game elements, game

mechanics, game dynamics and game techniques. The first three of these concepts have been provided by Werbach and Hunter (2012) whereas the fourth one (game techniques) is used by Chou (2015).

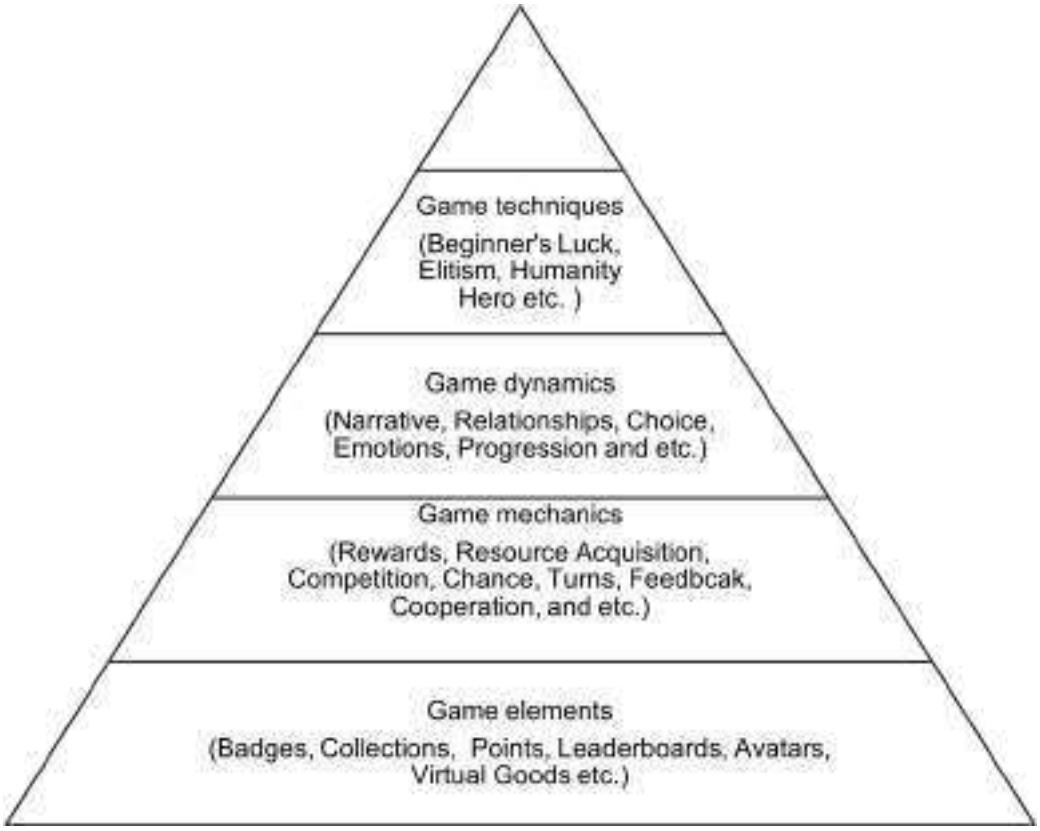


Figure 1. The relationship between game elements, game mechanics, game dynamics and game techniques (Chou, 2015; Werbach & Hunter, 2012).

However, the use of game elements or game mechanics is not free from some criticism. For example, Reeves and Read (2009) criticize their use and point out that it is possible to see some the so-called game elements such as avatars, ranks, levels or time pressure in other non-game contexts, identifying them as game specific elements or gameful may not be possible. Another problem with the definition of game elements is that they may not be perceived by different individuals in the same way. That is, they are subjective elements which can be perceived differently according to aims of designers. Moreover, Hanus and Fox (2015) criticize that how effective different game elements can be has not been tested enough. So, generalizing the findings of studies which investigate the effects of several game elements may not be possible.

When game mechanics are selected according to their needs and desires, it is possible to stimulate the behavior and motivate people. Some of these needs are rewards, status, accomplishment, self-expression, competition and altruism (Bunchball, 2010). Some people strive for getting rewards, tangible or intangible, after completing an action successfully. Points are mostly used as reward mechanisms in gamification. In addition, most people want to be recognized. Fame, prestige, attention, self-esteem and respect from others are important things for them and they engage in activities to meet these needs – to obtain a kind of status. In terms of accomplishment, it can be said that some people feel satisfied after accomplishing a challenging situation and attaining their goal. When the action provides optimum skill-challenge balance, they get motivated.

On the other hand, some people try to differentiate themselves from others by their autonomy and create their own identity. In addition, some people also enjoy to be in competitive environments. Of course, these are not the whole list and it cannot be said that an individual can become motivated just to meet one need. Instead, several needs can act together. So, it is important to take these needs into consideration before designing a gamification system. This may facilitate the process and increase the possibility of success. This is similar to what Chou (2015) tries to achieve with his 'Octalysis' framework. He states that beyond PBLs, different kinds of game mechanics can be applied while designing gamified systems according to what he calls eight CDs.

Why Gamification. Today the concept of gamification is being applied in many areas such as business, marketing, corporate management, wellness, education and so on. Moreover, Schell (2010) hypothetically envisions that in the future everything will be gamified because as Zichermann and Linder (2010) put forward many traditional non-game contexts including educational ones are not interesting. So, since games are fun, utilizing game-like features may help make those contexts be more attractive. Mainly what is aimed with gamification is to increase the engagement of individuals by improving their motivation levels (Jenkins, 2015). Deterding et al. (2011) state that game design elements do not only affect the engagement but also influence loyalty, fun and revenue. So, it can be said that basically, by using gamification it is aimed to improve individuals' motivation and engagement through game thinking and game design elements (Dicheva, et

al., 2015) and create positive behavioral change (Rapp, 2017; Seaborn & Fels, 2015).

Many studies provide enough rationale to use gamification. That is, many studies argue that positive outcomes will emerge when gamification is applied appropriately. For example, Jung et al. (2010) and Mekler, et al., (2013) state that gamification is successful in terms of obtaining desired outcomes by employing game elements in non-game contexts. Likewise, Kuo and Chuang (2016), who empirically investigate the application of gamification in online context of academic promotion and dissemination, suggest that gamification can be used to create attractive and motivating environments and can engage and retain users. They also point out that gamification can have positive impact.

When people play games, they can spend countless hours and try to develop their problem-solving skills (Gee, 2008). They not only enjoy themselves while playing but also develop their persistence, creativity and resilience by extended practices (McGonigal, 2011). Lee and Hammer (2011) state that gamification can enable the transfer of motivational power of games to the contexts in which motivational problems emerge such as school environments. However, they argue that before applying gamification, the theoretical rationale behind it should be clearly understood. In addition, what is meant by gamification should be clearly stated. The benefits and drawbacks of gamification, its sample implementations and future possibilities should be identified. Only after these are realized could successful interventions by using gamification be possible.

There may be several uses of gamification. Hanus and Fox (2015) point out that it can be used to tell narratives to change the context. In addition, by using gamification, it is possible to establish a competitive environment for students. It is also possible to encourage students to do certain actions by applying points and reward systems. In their longitudinal study in which they assess the effectiveness of gamification elements in terms of motivational, psychological and behavioral measures, they focus on the use of leaderboards, badges and competition as common game mechanics. They reach the conclusion that when these game mechanics are applied in appropriate way with clear learning goals, students may engage in meaningful activities that are gamified. Different game elements should be investigated in terms of their effectiveness. Also, the theoretical background

should be examined accordingly. In this way, optimum efficiency from gamification in classroom settings can be obtained.

However, game elements should not be limited to PBLs. According to Gåsland (2011), a lot of things that are found in games can be applied to non-game contexts as game mechanics. In terms of the application of gamification in non-game contexts, Deterding et al. (2011) state that the idea of using gamification should be taken into consideration in three separate categories which are meaning, mastery and autonomy. As Kim (2011) points out, for meaningful experiences, the system should allow the personalization and be designed in a way to promote positive emotions. In terms of mastery, individuals should be able to feel that they achieve interesting and challenging tasks. In turn, the positive feedback may encourage them and improve their feeling of accomplishment. But maybe one of the most important things is to be autonomous. Students should be able to take the responsibilities of their own learning experiences. Because when individuals become autonomous then intrinsic motivation is more easily achieved and there will be no need to force individuals to take action (Giannetto et al., 2013). From a similar point of view, Chou (2015) argues that one of the most important things is to integrate creativity CD into designs so that people get motivated because of the enjoyment of creating something. This means that people will be intrinsically motivated and this in turn will cause them to be more autonomous. There are other researchers who investigate the use of gamification in relation to its theoretical foundations. For example, Sailer, Hense, Mandl and Klevers (2013) mention six principal perspectives that are related to the use of gamification to enhance motivation: trait, behavioristic learning, cognitive, self-determination, interest and emotion (pp.31-33). In this vein, it could be said that utilizing game elements that are extrinsically rewarding may not be enough to reach optimum potential of gamification. It is highly important to understand the theoretical background of gamification to get the optimum benefit.

Educational Gamification

Using motivational interventions in various formats have long had a place in educational environments. Researchers have tried to improve the educational processes and outcomes in many ways. As one of these, teachers have been using

games to create challenging activities. According to Kapp (2012), in educational contexts games have already had considerable places in teaching and learning practices. In terms of education, Ellis, Heppell, Kirriemuir, Krotoski and McFarlane (2006) point out that games can be used to enable students to develop their knowledge and engage into the activities.

Games in education. As in other areas, in educational contexts, games have a considerable place. Gee (2014) states that games are frequently employed in education for meaningful experiences and social interactions. In recent years together with the developments in technology, the relationship between games and technology has received considerable attention. And researchers prepare annual reports about the emerging technologies in teaching and learning and try to predict when these technologies are likely to have a large impact on teaching and learning. According to a recently released NMC-Horizon (New Media Consortium) report by Jhonson, Becker, Estrada, and Freeman (2014), games and gamification are expected to influence educational technologies within two or three years then. Today, pervasive use of game applications such as Kahoot, Duolingo or Busuu and so on show that they are not wrong. In addition, this report indicates that when good games are applied to educational context, this may have positive impact on productivity and creativity levels of individuals.

There are studies whose results indicate positive effects of games in educational contexts (Di Bitonto, 2014; Gee, 2003; Gee, 2014; Michael and Chen, 2006; & Prensky, 2001). According to Michael and Chen (2006), it is possible to create a teaching tool by using games with educational techniques in such a way as to accomplish fun and serious goals at the same time. Similarly, Di Bitonto et al. (2014) states that gamified applications in education increase engagement and as a result can contribute to learning and teaching processes. Moreover, Prensky (2001) and Gee (2003) point out that video games have some potentials in learning environments. In literature, there are some different concepts that are used to define the application of games to educational contexts such as game-based learning, simulations, or serious games.

Although these all are related to games and game-like thinking in a way, there are some differences among their applications. When games are directly used for educational purposes, it is called game-based learning (Caponetto, Earp & Ott,

2014). In other words, when games are used at schools to improve learning experience, then it is called as game-based learning. There are numerous studies focusing on different aspects of using games in education. Becker (2021) points out that when a game is used intentionally as a tool or resource for learning, it is called an educational game. Some relevant works and literature reviews can be found in De Freitas (2006), Habgood (2007) and Klopfer, Osterweil, and Salen (2009). Another type of application is the simulations which are designed for educational purposes (Kapp, 2012). They are usually designed to allow individuals to practice their skills and knowledge. As with serious games, they are a special kind of games which do not aim to entertain. That is, as Ulicsak and Wright (2010) point out, when the main goal of using video games is learning itself, it is called as serious games. For example, games which are used to prevent world hunger (McGonigal, 2011) or the game called Global Conflicts aiming to help students learn about different conflicts in the world (Simoes et al., 2013) are among this kind of games.

In practice, integration of games in educational contexts can vary according to the approach, the contexts or the purpose chosen by teacher. According to Van Eck (2006), games can be used in education in three ways. Commercial off-the-shelf videogames can be used. The ready to use games developed by specialists and available commercially can be bought and their pre-determined contents can be used for educational purposes. As another approach, serious games can be utilized, which means that the main goal is not to entertain but to learn. Finally, students can create their own games. In this case, games can be created according to the content and the aim of the class.

Drawbacks of games. Although it can be said that games have various beneficial elements and game thinking has potential, trying to make use of these by using games themselves can cause several problems. Correspondingly, there are some problems related to game-based learning. First of all, as Johnson, Becker, Estrada and Freeman (2014) point out, creating educational video games is a costly process. In addition, as Simoes et al. (2013) argue, it is difficult to match the content of the games with the aims of education. Sometimes it is possible to come up with inappropriate or inaccurate content or the content games offer may not be complete. Simoes et al. (2013) also point out that if games which are developed by students

are to be applied, this process requires teachers to be competent in game development and game design, which is not usually possible.

Defining gamification of education. In 2010, the invention of the term gamification (Deterding et al., 2011) ignited the gamification research in many fields. As in other fields, researchers who are seeking better ways to apply games have developed the concept of gamification of education. Since then researchers have been in search of finding answers related to various dimensions of gamification. Moreover, the growing popularity of gamification as a research area was also approved by Gartner Institute's 2013 Hype Cycle which recognized the potential of gamification and placed it among the other promising technologies that were expected to reach their productivity plateau in five to ten years.

This report also showed that the use of gamification in educational contexts is getting popular each day. Because as Simoes et al. (2013) state, gamification allows the application of elements that really matter from the video games into education without using any specific game. What is aimed is to extract the elements that make games fun and enjoyable and integrate them into learning processes. In this way, students can learn not by playing specific games but they learn as if they play a game. Buckingham (2014) states that gamification can be utilized as a tool to increase individuals' motivation levels by integrating game elements such as the use of badges. In this way, it is possible to use the badges given to students for formative assessment purposes. Similarly, Glover, Campell, Latif, Norris, Toner and Tse (2012) acknowledge that the use of badges or other awarding systems could motivate students.

It is possible to find several definitions of gamification of education in literature. Actually, there are two approaches to defining gamification of education. In one of these, as it is stated before, gamification is the application of game elements and game design techniques to non-game contexts (Werbach & Hunter, 2012). Giannetto et al. (2013) state any classrooms can be accepted as non-game context since within learning environments the primary aim is not the entertainment itself. Thus, it may not be wrong to state that Werbach and Hunter's definition is a viable one. In another definition which focuses on the integration of game elements only Bellotti, Berta, De Gloria, Lavagnino, Dagnino, Antonaci and Ott (2013) see the application of gamification into educational context as the insertion of game

elements in learning contexts in design processes. In another study, Simoes et al. (2013) define gamification of education as “the use of game elements in a learning environment, usually with the support of ICT” (p. 347) by emphasizing the role of computers and information technologies in gamification applications.

On the other hand, there are researchers (Kapp, 2012; Seaborn & Fels, 2015) who do not differentiate gamification from other types of games, especially serious games. For example, Seaborn and Fels (2015) interpret gamification in this vein and argue that in the field of education, digital game-based learning and serious games are generally used to refer to gamification. Similarly, Kapp (2012) states that serious games are subsets of gamification and an educational content can be transformed into a game which is called a serious game through gamification process.

Why and how to gamify education. In literature, it is seen that gamification of education is mostly perceived positively because the present education systems are not flawless. Today’s schools face many problems in terms of motivating students because of their structures. Students are mostly unmotivated and the engagement levels are far from the desired levels (Lee & Hammer, 2011).

Thus, innovative solutions are needed to overcome problems that today’s’ schools have to face. In this respect, gamification can come to the fore and looks promising. Several researchers have reported positive views in terms of gamification of education. For example, Seixas et al. (2016) see gamification as a powerful engagement factor when applied according to the educational purposes. While designing systems with gamification, creating collaborative environments for students is significant. In a gamified system, a student should not be seen as a player who will be rewarded when he/she has completed an activity successfully. Instead, students should be seen as individuals who are responsible for their own knowledge development processes. Gamification can be used to facilitate these processes by turning them into funnier and more challenging ones. In their study, Seixas et al. (2016) show that gamification has positive impact on student engagement. Hanus and Fox (2015) state that with gamified classrooms, creating new ways of learning, which can motivate students more, and enabling them to enjoy themselves while doing tasks can be possible.

When different applications of gamification in educational environments are investigated, it is recognized that gamification can be applied in two ways. Mostly it has taken place in online educational environments rather than real classrooms. For example, codeacademy.com and khanacademy.org employ game elements to improve their users' engagement. It can be applied in online virtual environments and these can be used to support the education in physical environments. Or it can be applied in real classroom environments by changing the system. Caponetto, Earp and Ott (2014) also state that gamification techniques are mainly used to increase the motivation and engagement levels of students. Moreover, they point out that there are few studies that employ gamification in both face-to-face and e-learning environments. However, many of the studies are carried out online and interventions are made through e-learning platforms with gamified functions. However, according to Garvey (1990), applying gamification in physical classrooms is easier than applying it in digital environments because in real classrooms which are controllable environments for teachers, teachers can make use of various strategies while applying gamification.

On the other hand, in literature, there are studies that try to integrate several games to gamify the learning experience whereas there are others that try the integrate the game elements to the learning. That is, it can be said that there are two approaches to the question "Should we understand creating learning games or integrating game elements into the learning through gamification?" In terms of the educational contexts in which gamification is applied for, Dicheva et al. (2015) state that it is possible to group them under five categories: "gamifying courses without online support, gamifying MOOCs or online courses, gamifying blended learning courses, gamifying e-learning sites and developing gamification support platforms" (p. 81).

In addition, Dicheva et al. (2015) investigate the subject areas of the application of gamification. The categories they can identify are computer science, information technology, game programming, math/science/engineering and subject neutral areas. They conclude that although there are a good number of studies on gamification in education, the majority of them describe only some game mechanics and dynamics and try to repeat their use in educational context. In other words, in gamified educational environments, the same game elements are utilized without

producing new education-specific game elements. They think that if gamification design in educational environments is planned and applied properly, it can improve learning. In addition, future empirical research can help teachers decide which game elements to choose according to their own context to gamify their courses. In this respect, 'Octalysis' can be used to provide data about the weak and strong sides of courses. Teachers then could decide in which way to intervene and which game elements to employ.

Another perspective that should be taken into consideration while gamifying education is considering the level of application which can be at micro or macro levels. As Corbett (2010) says, in school environments gamification can be applied at micro level. That is, an individual teacher may gamify his/her class structure. At macro level, the whole school system or curricula can be gamified with the help of some gamification experts. Lee and Hammer (2011) take another approach. In Teachers College Colombia University, they create a modular toolkit for instructors and try to incorporate many gamified activities in the education. By using their toolkit, instructors could identify their own instructional educational needs and make necessary adaptations. Lee and Hammer (2011) think that when applied correctly, gamification could yield beneficial results for schools. But people who are going to apply gamification should be aware of the problems in the system which need to be fixed. They should design a new system by taking those problems into consideration. Having started the process, it should be checked frequently to see whether the interventions work or not.

Another issue to think about while gamifying education is to pay attention to the desired purpose. According to Lee and Hammer (2011) gamification interventions could be done on three main levels in educational environments: cognitive, emotional and social. On cognitive level, players try to overcome challenges and reach mastery in games (Koster, 2004). It is known that when the goals are adjusted according to the skill levels of the players, they motivate players (Chou, 2015). Through the subgoals, players can have the opportunity to determine they own routes. The existence of these subgoals also motivates them. This understanding may have a place in school environments. On their way to mastery, if students are aware of the path and determine their own goals, which are self-

concordant goals (Dörnyei et al., 2016), the end-goal which seems vague and unreachable at first may become reachable.

On emotional level, it is known that in games individuals experience a range of different emotions from curiosity to frustration (Lazzaro, 2004). One of the biggest challenges in school environments is the learned helplessness or the fear of failure. But in games, individuals are free to fail and with immediate feedback they can develop their skills. Lee and Hammer (2011) state that games can enable individuals establish positive relationships with games and develop resilience. They can try as many times as they can to become successful and with immediate feedback each time they fix a point and improve themselves. However, at schools, failure is not so welcomed and long feedback intervals cause problems. When gamification is applied in educational settings, failure can become a part of the process and by giving frequent feedback, students may find the chance to assess their own learning. Finally, on social level, in games, the social-interactional platforms may enable students to develop their social relationships and be recognized by others.

In terms of the game principles employed in gamification of education, Dicheva et al., (2015) found that game design principles are frequently used in educational contexts. Among these, “visual status, social engagement, freedom of choice, freedom to fail and rapid feedback are the most used ones” (p.79). Similarly, Barata et al. (2017) suggest that in a gamified educational system, there should be options for students to allow them to learn from their failures by trial and decide their own learning paths. In addition, they state that as students are rewarded for their achievements, the experience itself becomes more competitive. Since over-competitive environments may result in disengagement, new occasions that give new chances to students to collaborate should be provided and thus the competitive balance should be established. As in a game, students can exploit the rules to be more successful than their peers by focusing on the grades. So, performance assessment tools focusing on quality should be adapted.

As in other fields in which gamification is applied, for gamification of education, as Lister and College (2015) state, the most frequently preferred game elements are points, badges, leaderboards (PBLs), levels, progress bars. Levels as game elements are used to divide a game into small and more attainable separate pieces so that students can strive to achieve the next goal and try to level up

(Gåsland, 2011). In addition, there are studies which investigate the role of such specific game elements in gamified educational contexts. For example, in their study which provides empirical evidence of the impact of game mechanics in an education-related research course in Asian context, Hew et al. (2016) point out that game mechanics such as PBLs act as powerful incentives and do not have negative effects on students' learning of factual knowledge. When a student completes a task successfully, badges for achievements can be given as symbolic awards (Abramovich, Schunn, Higashi, 2013). But as Abramovich et al. (2013) argue, learners from different levels can give different reactions toward badges. Students who show low performance demand badges due to their participation whereas high-performing students demand badges because of acquiring a skill. Deci, Koestner and Ryan (1999) point out that using badges and reward systems may negatively impact student motivation and learning. Hanus and Fox (2015), who investigate the effectiveness of certain game elements which can be applied by a traditional teacher in a traditional classroom environment, state that using some game mechanics such as leaderboards, badges and competition may have a harmful effect on motivation, satisfaction, student empowerment and learning outcomes. Extrinsic rewards such as badges can work when an individual finds a class boring. However, for an individual who already finds the class interesting and thus extrinsic rewards may actually harm their intrinsic motivation.

Thus, it can be said that use of extrinsic reward mechanisms and game elements yield different results. And researchers have different point of views regarding the effects of them. For example Goehle. (2013) say that by using PBLs, a competitive environment can be created and students may feel sense of belonging to a group and they can compare their performances to those of their peers. Similarly, there are various studies that show positive impact (e.g. Barata, Gama, Jorge and Gonçalves, 2013; Charles, Charles, McNeill, Bustard, & Black, 2011; De-Marcos et al., 2014; Goehle, 2013). However, there are also some studies that find them demotivating (Berkling & Thomas, 2013; and Hanus & Fox, 2015). So, as Seixas et al., (2016) state, gamified applications in education should not be limited to PBLs. Various gamification strategies should be used to create environments in which students' creative abilities are uncovered, errors are not at the center,

students can collaboratively work together and can make free choices according to their needs.

However, these elements have the risk of being superficial when applied inappropriately. As Stott and Neustaedter (2013) point out, apart from integrating PBLs kind of game mechanics into gamification of learning environments, game designs should also include some other features. There should be instant feedback mechanisms. In addition, in gamified learning environments, students should feel free to fail. Progression and narrative should be included into the designs and so on.

It is possible to see some kinds of game mechanics in today's schools (Lee & Hammer, 2011). Students get points for their successes and they are graded. In addition, they are grouped according to their levels which means that they can level up or down. When these are taken into consideration, it can be thought that schools are already gamified. However, in reality, despite those game elements, it is almost impossible to talk about successful gamified experiences because of the low engagement levels, cheating, dissatisfaction, learned helplessness observed at schools. In this respect, Lee and Hammer (2011) conclude that the existence of game-like elements does not guarantee successful gamification. So "under what circumstances game elements can drive learning behavior" (p. 2) should be taken into consideration.

In the field of education, which is the most investigated field in terms of gamification, it is possible to talk about many other studies as examples (e.g. Gåsland, 2011; Foster, Sheridan, Irish, & Frost, 2012; Li, Grossman, & Fitzmaurice, 2012; Denny, 2013; Dominguez, Saenz-de-Navarrete, de-Marcos, Fernández-Sanz, Pagés, Martínez-Herráiz, 2013 & Goehle, 2013). For example, Giannetto et al. (2013) develop a gamified system for social learning environments in higher education lecture classrooms. They aim at improving the engagement levels of students by using their system which encourages five types of behaviors: social, intelligent, helpful, inquisitive and hardworking (p.202). They also state that applying gamified systems to non-game environments will have a positive impact on students' lives if it is carried out appropriately.

There are studies which focus on the role of gamification in real classroom environments. In one of these, Sheldon (2012) has tried to increase students' engagement levels and make classes more fun and interesting with gamification. What is important in his design is that the learning experience is designed without using technology. However, statistical evidence which shows that this kind of approach could work is missing. In another one, Barata et al. (2013) designed a gamified course to see how effectively gamification can be utilized for better student engagement rates. In terms of lecture attendance, they could not find any difference. But they see a significant increase in students' engagement rates by using their design. In addition, they find out that better scores were received from the gamified learning experience and the grade gap between successful and less successful students decreased.

Recent studies have begun to focus on the role of individual differences instead of the role of specific game elements. Barata et al. (2017) design a study through which they can understand what different behaviors may emerge in a gamified learning and how different students are engaged by the game. They state that their gamified version of the course results in increased students' engagement and students find the course more interesting and motivating when they compare it to the traditional one. At the end of their three-year study, they find out that there are different types of students with different traits. They identify four types of students which are achievers, regular, halfhearted, and underachievers. According to them, achieving is highly important for achievers and they try to collect all the experience points (XP) they could get. They called the students who are above the average in terms of their performance as regular students. These kinds of students try to balance their achievements in game with more traditional evaluation mechanisms. Halfhearted students perform below the average and they seem to disregard some course components. Finally, underachiever students have the lowest performances and strive for just to get the passing grade.

In addition, it is possible to find some studies which provide frameworks for gamification of education (Simoes, et al., 2013). In one of these studies, Simoes et al. (2013) investigate how to apply certain social game elements to social learning environments. They propose a framework to gamify a virtual learning environment

and offer a step-by-step guide in designing socially gamified learning contents. However, there is no evidence whether their framework works or not.

Gamification in EFL context. When literature on gamification in English as a second language classrooms is investigated, it is seen that there are few studies although it is possible to see relatively more studies carried out in other areas of education. However, this does not mean that in second language (L2) learning contexts, it is not possible to allow teachers to redesign the end-goal by gamifying the processes. In one of these few studies, Baber (2015) investigates the distribution of gamification studies for the five-year period both in general education and in English as a Foreign Language education. Interestingly, his study reveals that although the number of general gamification studies is over thousands, there are only nine studies which directly deal with the "gamification TEFL". In his study, Baber (2015) offers two types of gamified physical EFL classes. In these classes, students should complete the predetermined tasks to level up. Both systems include several types of game elements such as points, badges and levels and take the basic game principles into consideration such as immediate feedback, challenge, motivation and engagement. However, Baber does not mention any kind of statistical findings whether these work out or not.

Therefore, it can be said that for language learning, gamification is usually applied in mobile learning applications. There are many applications that utilize game elements such as Memrise, LectureQuiz, Duolingo, Lingobee and Quizfun (Berg, 2013). These kinds of mobile applications, which are created for language learning, are among the most popular gamification examples.

Gamification in Turkish context. In Turkish context, there are several studies investigating gamification (Alsancak Sırakaya, 2017; Birant, 2014; Bozkurt & Genç-Kumtepe, 2014; Çağlar & Kocadere, 2015; Erdoğdu & Karataş, 2016; Gökkaya, 2014; Güler & Güler, 2015; Karataş, 2014; Sarı & Altun 2016; Şahin & Samur, 2017; Tunga & İnceoğlu, 2016; Yıldırım & Demir, 2014, 2015; Yıldırım, 2016). In one of these studies, Şahin and Samur (2017) conduct a literature review in Turkish context. Their study reveals that mostly self-determination theory, flow theory, Fogg behavior model and dynamics-mechanics-components approach are theoretically associated with gamification. In addition, in Turkish context, most of the

studies deal with gamification in the scope of motivation and PBLs and feedback and rewards are the most used game elements.

Many studies in Turkey investigate gamification application in terms its effect on motivation. Yıldırım and Demir (2014) state that game designs aiming at increasing motivation may also increase students' engagement. In another study, Güler and Güler (2015) point out that integrating game elements into the education designs can positively affect students' motivation. Similarly Kocadere and Çağlar (2015), who conduct a study in which they designed a gamified assessment system, state that their system has positive effect on students' motivation, enjoyment and learning. They also state that gamified assessment system causes student experience flow and the system reduces their anxiety.

In another study conducted in Turkey, Turan, et al., (2016) investigate the effects of gamification on students' cognitive load levels, achievements and perceptions by comparing gamification strategies to traditional methods in an online learning environment. Their findings point out that gamification can increase students' both cognitive load and achievement levels and students have positive thoughts about gamified learning experiences. There are also completed dissertations (Erümit, 2016; Meşe, 2016; Yıldırım, 2016; and Ersoy, 2017) and theses in Turkish context.

Drawbacks of gamification of education. Although the integration of gamification into the educational environments is generally found beneficial by researchers, it is not free from some drawbacks. Dicheva et al., (2015) state that PBLs should not be considered as the result of gamifying education since they have been in use in educational context for many years. Lee and Hammer (2011) point out some of these. First, gamifying a learning activity or class may increase the workload of the teacher. In addition, using too many external rewards or motivational elements may result in reverse effect. Another issue is that, one of the main pillars of games is the freedom of choice, failure or control. When gamified systems become mandatory and students are forced to take part in games involuntarily, then it means that there will be no difference between traditional school experiences and the gamified experiences. To avoid such situations, Lee and Hammer (2011) suggest that gamification should be carefully planned and should be based on

existing research. In addition, the process should be assessed frequently to see whether it works or not.

There are researchers who advocate the use of gamification on digital platforms. Dominguez et al. (2013) state that it is possible to increase students' motivation levels by using gamification in e-learning platforms. Their qualitative analysis shows that although gamification can emotionally and socially effect students' motivation, this finding is not conclusive since there are some students who do not find the gamified system fun at all. They point out that by focusing on only extrinsic rewards, desired results from gamification cannot be obtained. Their quantitative data indicate that on cognitive level, gamification do not have any important effect on students' motivation. According to their findings, some of the students who have followed the traditional classes do better than the ones who have attended gamified program.

As Lister and College (2015) state, while designing gamified educational systems, designers not only pay attention to the utilization of PBLs but also take individual differences and profiles of learners into consideration so as not to produce negative effect with gamification. There are also other problems in terms of gamification designs apart from these differences. To get better results, as Dominguez et al., (2013) point out, a good gamification design requires large amount of investment and set-up. Moreover, technical issues and usability can cause problems.

Drawbacks of gamification. Some of the biggest criticisms towards gamification are about its definition. Seaborn and Fels (2015) state that there is a lack of consensus in terms of standard definition of gamification. Another criticism is that many studies (87%) do not address any theoretical foundations (Seaborn & Fels, 2015). They argue that in many studies the theoretical foundations are inconsistently referenced and interpreted. Moreover, there is a gap between theory and practice and more empirical studies are needed.

The findings and results given in studies are also under criticism because as Hamari et al., (2014) mention, in many studies gamification yield positive results/effects. However, findings from most of the quantitative studies which mention positive effect of gamification cannot be generalized because it is not

possible to talk about the real effects of gamification by just examining the perceptions of participants. On the other hand, when the qualitative studies are investigated, they say that the diverse nature of gamification is not examined enough.

Another criticism of gamification is that the behavioral changes may not be the effect of using certain types of game mechanics, but because of the feeling of novelty and the curiosity created with those game mechanics. This is called as novelty theory of gamification (Farzan, DiMicco, Millen, Dugan, Geyer, & Brownholtz, 2008; Koivisto & Hamari, 2014). In the same vein, Hamari (2013) and Farzan et al. (2008) state that gamification can only lead to short-term results, which means that the motivational and behavioral effects of gamification may be temporary. After the participants get used to the gamified system and the novelty has faded away, the motivational boost may be lost.

From another point of view, some researchers debate the dominance of digital world in gamification applications. According to Chang (2012), in most of the studies on gamification, virtual things are given so much importance that it is almost impossible to realize the claims of those who favor it. Also, Deterding et al. (2011) state that although most of the gamification examples are digital, there is no need to restrict it within digital technology. Since in most studies game mechanics that can be found in video games are investigated, this causes people think that gamification is only related to video games and only programmers can do it. However, as Schell (2008) states, people do not need computers to play many games such as board games, card games or athletic games. Actually, he points out that video games are just a new version of traditional games offered from new channels.

When these traditional games are examined, it is seen that as Chou (2015) argues, they have passed the time test and they are still being played by many people. They have been successful for ages without the contribution of newest technology. Thus, as Schell (2008) points out, it is possible to identify some common underlying psychological principles that most kinds of games share without looking at where the games are played. If these common principles can be identified, then it may become possible to integrate them in other systems. That is, instead of using just video game elements for gamification, it is possible to find other game elements

which are appropriately used both in real and virtual gamification environments. And to do this, a person does not need to be a computer engineer or programmer.

In terms of the problems related to game mechanics, another analysis is made by Hanus and Fox (2015) who state that there are so many game mechanics which can be applied in countless combinations and ways. Therefore, it may not be possible to study each of them. In this respect, they suggest that instead of focusing on the diversified game elements and mechanics in studies, investigating the effectiveness of drives underlying them may be more helpful. Similarly, Antin (2012) states that extrinsic and superficial rewards like PBLs are not the main driving factors behind the real gamification. Instead, rewards related to social factors such as self-efficacy, community and peer approval enable users to continue their involvement. Relatedly, in most of the commercial applications and academic research, only PBLs are employed as game elements. The use of PBLs as game elements is criticized by Robertson (2010). She states that although PBLs are not so important elements of games they are used as if they were the core units of games. Bogost (2011) shares a similar view and states that using extrinsic motivation too much ruins the charm of games. This shows, as Rapp (2017) states, that the ways of using games to improve human-computer interaction and interactive systems are not sufficiently investigated and there is a need to go beyond current gamification practices.

In terms of the motivational effects of gamification, there are researchers who disagree with the positivity of gamification. According to Rapp (2017), although it is generally accepted that gamification is efficient in motivating users, there are studies stating that gamification has some limits and there is fuzziness about how to redesign gamification strategies (Nicholson, 2012; and Rao, 2013). In addition, Deci, et al. (1999) point out that using badges and reward systems may negatively impact student motivation and learning.

These kinds of criticisms can also be found in educational gamification. In one of the studies, Hanus and Fox (2015) investigate the effectiveness of certain game elements which can be applied by a traditional teacher in a traditional classroom environment. They state that using some game mechanics such as leaderboards, badges and competition may have a harmful effect on motivation, satisfaction, student empowerment and learning outcomes. Extrinsic rewards such

as badges can work when an individual finds a class boring. However, for an individual who is already finds the class interesting and thus extrinsic rewards may harm their intrinsic motivation. With a similar point of view, Deterding (2013) argues that existing systems have some problems because of the superficial application of single game elements in environments. In other words, when only the extrinsic rewards which cannot foster intrinsic motivation are aimed, failure in terms of promoting meaningful experiences may become inevitable.

Criticisms related to the application of gamification are not limited to above mentioned points. From another perspective, Landers and Armstrong (2017) demonstrate users' attitudes and previous technological experiences play a significant role in gamification. They say that when users have less game experience and poorer attitudes towards games, they may not benefit from the gamified instructions as much as others. That is, the desired instructional outcomes may not be obtained if participants' attitudes towards game-based learning and experience with video games are low.

Another criticism is made by Fitz-Water, Jhonson, Wyeth and Tjondronegoro (2017). According to them, designing gamification is not an easy task since the context and users may affect design process. In their study, they find out that although using game elements can positively affect participants' enjoyment, motivation and engagement, this does not necessarily mean that this desired behavioral change will occur. That is, behavioral change may not always be obtained whenever game elements are added into non-game contexts.

In addition to these, some researchers highlight the importance of when or when not to apply gamification. Since most studies do not put reasonable rationales for applying gamification, it may not be possible to decide whether it is necessary or not. Landers and Callan (2011) argue that while designing training with gamification, training program needs, specified program objectives and a valid reason to apply gamified designs are crucial because whereas sometimes gamification is the right tool, some other times it may not be. Thus, applying gamification just because it is available may not yield desired results. Also, as Landers and Armstrong (2017) suggest, trainee characteristic such as attitudes, previous experiences in terms of technology and video games need to be carefully taken into consideration before applying gamification.

In short, it can be said that studies conducted so far are not free from some drawbacks. Some studies have theoretical problems whereas others have problems in terms of the generalizability of their findings. In addition, the excessive use of game elements and mechanics from video games may cause problems by ignoring game elements and mechanics specific to non-virtual games. Moreover, many studies focus on just the use of PBLs or similar game elements instead of focusing on the drives. This thwarts the main aim of gamification.

Gamification Models and Frameworks.

When the literature so far is inspected, it is understood that modifying educational contexts by using gamification can yield positive outcomes. However, as Stott and Neustaedter (2013) highlight, the way of applying gamification is important for successful applications. Because applying game elements merely may not be enough. Researchers like Lee and Hammer (2011) and Dicheva et al., (2015) emphasize that to create successful gamification systems, it is important to make appropriate plans. In this respect, several researchers (Aleven, et al., 2010; Aparicio et al., 2012; Chou, 2015; Nicholson, 2012; Rapp, 2017; Sakamoto et al., 2012) provide frameworks for gamification interventions in educational environments. According to Huang and Soman (2013), frameworks can enable teachers to re-plan their classes. Huang and Soman (2013) offer a five-step model for this purpose. Accordingly, first educators need to know their students and they need to be aware of different needs of different students. Then as the second step they should define their learning objectives. They should define their goals both as general and specific ones. Afterwards, they need to structure the learning experience by identifying the main points. Then the resources should be identified (pp.7-13).

One of the frameworks with theoretical foundations is developed by Aparicio et al. (2012). Their framework is based on self-determination theory. That is, autonomy, competence and social relatedness (Ryan & Deci, 2000) are the main concepts. Identification of the main objective, identification of transversal objective, determining the game mechanics to be employed and finally the evaluation of the framework in applied systems are the four parts of their framework. Another framework which focuses on motivational factors is made by Song and Zhang (2008). Song and Zhang (2008) proposes another design in which there are ten

design principles that aim to meet individual needs in five motivation sources: 1. Psychological (autonomy and self), 2. Cognitive (competence and achievement), 3. Social, psychological (relatedness), 4. Social, psychological (power, leadership and followership and 5. Emotional (emotion and affect) (p.?). Song and Zhang (2008) and Aparicio et al.'s (2012) models propose that it will be more appropriate to choose game elements or mechanics according to the psychological drives that have impact on students' motivation.

There are other frameworks offered by different researchers. For example, Nicholson (2012) purposes a user-centered framework for meaningful gamification Accordingly, in this framework intrinsic and internal motivation is focused rather than extrinsic or external motivation and at the core of the framework is the 'organismic integration theory', a theory derived from SDT, explains how motivation can intentionally can be mediated by internal and external methods of control.

Another framework is proposed by Sakamoto et al. (2012). Their value-based gamification framework aims at encouraging and harnessing intrinsic motivation. There are five values in the framework: information as in prompt and necessary; empathic values, based on virtual characters and social engagement; persuasive values, a form of information that provides an outlook based on current behaviors, actions and outcomes; economic values, related to collection and ownership; and ideological values, defined as beliefs implicitly supported through stories or message formats (p.423).

As a result of his ethnographic study in which he aims to find new meaningful game elements, Rapp (2017) identifies nine recommendations to enhance user engagement and increase motivation for participation. These include identification and empathy together with past and future selves, rewards, social presence, self-organization, cooperation and friendship, competition, freedom and journey (p. 465). According to Rapp (2017), these recommendations help to improve intrinsic motivation because they employ cognitive, social and emotional drives instead of merely using extrinsic rewards which results in mechanic behaviors. In addition, these recommendations allow going beyond PBLs since they provide new and different kinds of game elements. Finally, it becomes possible to study systematic design strategies instead of applying game elements one-by-one. When Rapp's model is investigated, it is seen that recommendations aim at improving user's

digital images. By using different kinds of game elements such as character customization, role playing or character evolution, it is aimed to arouse user's memories and drive his/her behaviors.

Second recommendation aims to get rid of over-use of extrinsic rewards by providing meaningful and diversified rewards according to user's competence and values. Game elements that enable the competence-reward connection can be used for this purpose.

Recommendation three, which is social presence, aims at making use of user's social presence and increasing their participation and performances. Forming pick-up groups, add-ons for exposing player's stats and leave party can be counted as some examples of game elements that can be used.

In terms of recommendation four, different social structures should be created in a social environment to be able to meet different individual differences. Game elements which are given as examples are self-organization guilds and proliferation of different guild types.

Recommendation five (a) aims at improving cooperation among groups and making use of friendship among individuals for long-lasting effects. Guild-charts and forums, common aims, private chats, whispering of newsfeed can be used as game elements.

Recommendation five (b) suggests improving inter-group comparison without supporting conflicting interests while designing for competition. In this way, users can be encouraged to work for improving their own groups by cooperating and participating more. As game elements, challenge against the game or indication of character's association to one guild, can be utilized.

Recommendations six (a) and six (b) suggest not using boring, easy and certain aspects of games too much. These recommendations suggest that user's intrinsic motivation can be increased by providing diversified and rich experiences and allowing users to make their own choices freely to progress. Freedom of exploration, overabundance of quests, professions and achievements, experience points at first; then items, instances and friends, choices can be utilized as game elements according to these recommendations (p. 465).

Apart from these frameworks, another comprehensive framework is offered by Chou (2015) which is called as Octalysis. Chou argues that gamification is not limited to PBLs. What is important is to take the drives underlying games into consideration so that it can become possible to utilize any game element or game mechanic according to the circumstances. He also adds that his model can be applied to any kind of system at the center of which is human and employed for motivational optimization. Chou's (2015) model includes eight Core Drives (CDs) which are epic winning and meaning, development and accomplishment, empowerment of creativity and feedback, ownership and possession, social influence and relatedness, scarcity and impatience, unpredictability and curiosity, and loss and avoidance.

Rapp's (2017) recommendations and Chou's model resemble each other. Both of them take the psychological drives at the center and allow choosing appropriate game elements depending on these drives. However, while Rapp's (2017) nine recommendations offer new ways of designing more engaging and motivating systems, how to apply them into new fields is not certain. Chou's (2015) model can be used as a diagnostic tool and allows teachers to choose game elements according to their needs. In addition, Chou offers several game techniques to use for each drive and this allows teachers to fine tune their interventions. However, what is missing is the statistical evidence that proves that these eight CDs can work together. In the following part, Chou's (2015) Octalysis model is investigated in detail and its underlying theoretical structures are explored.

Octalysis

The starting point of Octalysis philosophy is to create a human focused approach. Chou (2015) sates that most of the existing systems are function-focused which focus on completing the job. He exemplifies this with workplace analogy. Accordingly, in function-focused workplace designs, the payment for your job should be enough to get you to complete your job successfully. Which is not the case. People should be motivated and the design of the workplace should be human-focused.

For Chou, gamification is at the core of human-focused designs. He defines it "the craft of deriving all the fun and addicting elements found in games and

applying them to real-world or productive activities" (p. 8). What is important is, instead of completely focusing on a fully efficient system, to optimize the system for human. And gamification is the first industry that has succeeded in enabling human-focused design. The reason to build a design framework on gamification is that this industry has spent years to find the ways to motivate and enable engagement. In addition, the power of games and game elements in motivating people is another reason that attracts Chou's attention. Chou (2015) states that "games have the amazing ability to keep people engaged for a long time, build relationships and trust between people, and develop their creative potentials" (p. 10). In games, people not only entertain themselves but they become productive and the more time they spend, the more productive they become. They build relationships, create value, and solve the hardest problems.

Human motivation is not so simple to define at first glance. However, Chou states that there are eight fundamental factors which he places on a unique shape beneath why someone wants to do something. He names these eight Core Drives (CD) as: 'epic meaning and calling', 'development and accomplishment', 'empowerment of creativity and feedback', 'ownership and possession', 'social influence and relatedness', 'scarcity and relatedness', 'unpredictability and curiosity', and 'loss and avoidance' (pp. 25-28). However, naming these eight CDs are not enough. To make it more useful and actionable, Chou (2015) created Octalysis. He states that Octalysis is an octagon where each point represents a CD. The CDs are purposefully placed so that it can be possible to create rhythm, symmetry, and meaningful systems.

As indicated before, one of the important facets of Octalysis is its ability to utilize different theoretical fundamentals. Chou (2015) not only names and is able to provide graphical representation of these CDs, he is also able to integrate one of the most influential motivational theories which is 'Self-Determination Theory' (Deci & Ryan, 2004) into the same design. For this, he uses the terms left and right brain distinction and positive and negative motivators distinction as he places these eight CDs on the Octalysis (p. 28). The left side of the Octalysis is related with logic, calculations, and ownership. Left Brain CDs (CD 2, CD4 and CD6) are extrinsic motivators – a user is motivated because they want to obtain something, whether it be an intangible good or a tangible item. On the other hand, the CDs on the right

side of the Octalysis (CD3, CD5 and CD7) consists of creativity, self-expression, and social aspects. Right Brain CDs are intrinsic motivators: a user doesn't need a goal or reward to use his/her creativity, to hang out with friends, or to feel the suspense of unpredictability – the activity itself is rewarding on its own (p.29).

Another major distinction used in Octalysis places 'Approach and Avoidance' dichotomy which is done by dividing the drives as positive (white hat) and negative (black hat) drives. On the top of Octalysis, there are positive motivators (CD1, CD2 and CD3). If something is engaging because it lets you express your creativity, makes you feel successful through skill mastery, and gives you a higher sense of meaning, it makes you feel very good and powerful. The CDs which are placed on the bottom (CD6, CD7 and CD8) are negative motivators. If you always do something because you don't know what will happen next, you are constantly in fear of losing something, or because there are things you can't have, even though you are extremely motivated to take action. The following example can be helpful to explain white and black hat drives: A student may work hard in a motivated way for an exam, but studying hard may be because of the pure enjoyment of learning and approaching a meaningful and bigger goal or may be because of just not to fail in the exam and become unsuccessful. In the following section, each CD is explained in a more detailed way.

Core Drives of Octalysis

CD1: Epic Meaning and Calling. People want to be or need to be a part of something big. When they manage this, they take the action for the sake of a greater thing. Chou (2015) explains this with Wikipedia example. People examine hundreds of pages and try to find out mistakes and correct them and, they try to keep the pages up to date and contribute by adding new information. And they do so without being paid anything since they think that their work is important and the thing they do affect many people around the world in a positive way, which means that they have an important mission.

Another example from Chou's own web page exemplifies how CD1 works (https://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/). During a show in which western and eastern parents encourage their children to get the scholarship after competing in math, science and language. In

this competition, westerner families motivate their children with CD2 (development and achievement) and CD 4 (ownership and possession). The important thing for them is accomplishment and in case of failure they do not bother too much and encourage children to try harder next time. On the other hand, easterner parents motivate children with CD1 (epic winning and calling). It means that wining the competition is related to bringing honor to the family. Here they motivate children to win not for themselves but for parents and family honor.

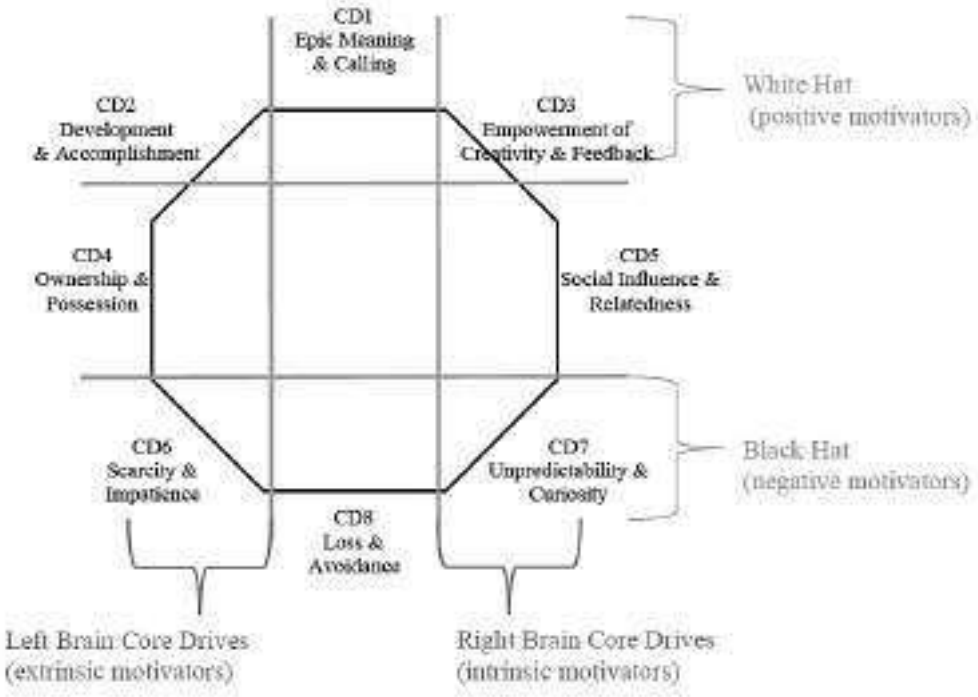


Figure 2. Structure of Octalysis (Chou, 2015, p. 23).

CD2: Development and Accomplishment. This is the drive that "motivates through a sense of growth towards a goal and accomplishing it" (Chou, 2015; p. 25) What makes primary school children cheer up after getting positive feedback from their teachers, or academicians share the visuals of their latest publications or presentations on social media can be given as examples to CD2. In this sense, it can be said that while people strive to overcome a challenge and try to achieve their goals, CD2 is what motivates them. So, to motivate users to go on playing, there needs to be challenges which are optimized. That means that they must be neither too difficult nor too easy. If there were not any rules and limitations which challenge people to overcome, basketball would be nothing but a game that require putting a ball into a hole which is not so funny and doesn't get people engaged.

CD3: Empowerment of Creativity and Feedback. CD3 is located on the top right side of the Octalysis which is called as the Golden Corner because according to Chou (2015), the top drives on Octalysis are called white hat gamification which “makes people feel good” (p. 26). Moreover, on the right side of the Octalysis, right brain gamification exists. In other words, as stated before, this part is about intrinsic motivation which “makes you feel good and you take action because you find the action enjoyable on its own” (p.25). For example, kids play with their dolls or toy cars. They just enjoy the activity for its own sake because it activates their creativity. They create their own garage or house by themselves and usually they do not need to overcome something. They try different combinations and scenarios every time and keep playing for hours. Chou (2015) states that CD3 is one of the hardest drives to implement.

CD4: Ownership and Possession. CD4 makes you feel the ownership of the experience. It is stated that when combined with CD8 (loss and avoidance), a person cannot risk losing everything. Youtube, Netflix or online shopping sites offer personalized content that is derived from user data and history every time you enter the system. As time passes, you cannot give up because you do not want to lose your personalized and customized data and recommendations. Sometimes it is seen that individuals have difficulty in throwing their unused belongings or books. Some companies offer games during their marketing. They ask you collect items some of which will be very few by buying their products and complete a task to get the reward which can be money or something big in exchange. Although these rewards worth nothing people continue to collect and try to complete the task in a highly motivated way.

CD5: Social Influence and Relatedness. CD5 puts emphasis on the effect of what other people think, do or say on your choice of activity. Chou states that this drive is “the engine behind themes like mentorship, competition, envy, group quests, social treasures and companionship” (p.27). The relatedness part of this drive is about the emotional attachments and the effect of nostalgia on feelings. It is highly possible to see people do business more easily if they know the other person who is from the same hometown. Or people tend to buy a product that reminds them their past more easily. You can easily observe students want to be together with their best friends or when they enter a foreign environment they prefer to stand by

their friends. That's why many companies try to ensure you invite your friends to create your own society which will have common historical and social ties that will make you be active.

CD6: Scarcity and Impatience*.* At the core of CD6 which is a black hat CD is the 'prestige' as Chou (2015) states. He gives the acceptance rates of the most prestigious universities as example and concludes that Harvard which has the acceptance rate of 5.2% would be worthless if it had 84% acceptance rate. Another example is from Oren Klaff's book Pitch Anything (2011). Accordingly, people want to have the things which they think is hard to get. And the things that are difficult to reach are perceived to be valuable. Chou mentions about a diamond company which controls all the diamond trade across the world. He says that diamonds are not so valuable because they are rare but because this company makes them scarce to increase their value.

In most mobile games to which we are addicted, designers embed 'torture breaks' to compel users to obsess and keep coming back to play. For example, in Candy Crush, you must wait for 25 minutes for another life after you spend your lives. In this way, since you have waited, you want to go back and play. In a classroom, when a teacher puts a candy as prize for the correct answer, children will engage in that task in a more motivated way than usual.

In this sense, it can be said that if we are unable to obtain something immediately or with great difficulty, then we become motivated. If something is around us and we can reach it easily, it becomes unimportant. Because we always need to get the out of reach one. In other words, we understand the value of something when we lose it. Or as Chou (2015) states, humans tend to search for what is scarce and unavailable. Also, if they realize that something is going to fade away soon and will not be available, they act toward getting it. This means that, if you want to increase a certain behavior, one powerful way is to place a limit on that activity. Of course, you don't necessarily want to limit the activity so much that you are losing key metrics from that limit. The best way is to first find an 'upper edge' of that metric which simply forms a perceived sense of scarcity, but doesn't necessarily limit the behavior (Chou, 2015; p.27).

CD 6 is the first source of generating CD 3: Empowerment of Creativity and Feedback in the system. When used with CD 7: Unpredictability and Curiosity, CD6 becomes a great engine to drive online consumer action. Finally, working alongside CD 8: Loss and Avoidance, CD6 becomes a powerful force that not only pushes for action, but pushes for action with extremely strong urgency (Chou, 2015; p.267).

CD7 - Unpredictability and Curiosity. CD7 is a black hat CD and intrinsically thrills us. As far as company sustains curiosity and unpredictability, it can make its customers stay longer with them and ensure engagement. However, CD7 may not be so innocent because it can also result in addiction when it is combined with CD6, another black hat CD. Although winning in a lottery or gambling game is far less than being crushed by a meteorite, people get obsessed into playing these kinds of games because of the thrill of chance which triggers the hopes to win.

When looked at from this view, it may seem that investing slow and steady way may be more beneficial. But people always feel excitement in almost everything and want to try their chance. Chou (2015) mentions a few companies which have used this CD and have achieved to sell their free samples for money. Indeed, free samples as the name indicates should be given out free. But integrating curiosity and make people curious about the product have enabled companies make money from those free samples. People want to pay for a box of free samples every month because they do not know what will be inside the packet. In addition, in one occasion a company combined CD5 and CD7 and achieved to sell more. You can see lots of videos on Youtube recorded by little children and that have thousands of followers. In these videos which are very simple and amateur, children open a giant egg which is full of usual toys that every child knows in front of camera. However, thousands of children like to follow these videos just because they want to find out what will come out of the egg.

In another example Chou mentions the Skinner Box experiment which was carried out by B. F. Skinner. In this experiment, Skinner puts animals into a box with a leveler and when the animal presses the leveler it gets food. When animals are given food for each time they press the button, their desire to continue ceases when they get full. However, when the food is supplied with unequal intervals, animal continue to press the button no matter how much it gets and whether it is not hungry

anymore. The thing is that if something triggers our intrinsic curiosity, we may sometimes forget about the extrinsic reward.

In short, as Chou states, while working on engagement design seeking out the ways of integrating randomness and chance into the system may improve the productivity. As far as there is unpredictability and curiosity, there will more engagement. CD7 can be combined with white or black hat CDs for better efficiency. However, it should not be forgotten that when accompanied with black hat CDs, it can cause obsessive behaviors.

CD8 - Loss and Avoidance. What lies behind CD8 is being afraid of losing something or undesirable events. If it can be tuned finely this black hat CD can be utilized in a very effective way. But in the long term, it can cause discomfort. We do not want to lose the things that we have invested and may sometimes be reluctant to change our behavior to preserve it even there is a better alternative. As Chou puts forward, there are many situations in our lives in which we do not want to lose our investment of time, money, effort and so on. To preserve our ego or sense of self, we may sometimes refuse to give up. The more one invests on something, the more it becomes harder to leave it as in the case of Farmville game. In this game, you invest time to grow your crops and you must log in the game in a given time to take care of your farm. If you miss the given time, your previous efforts go. So as time passes and the time and effort that people put increases they try to keep up and even get up in the middle of the night or at dawn for harvest.

We not only fear to lose what we have; we also fear to lose what we could have had. You can see many online shopping sites that has countdown clocks. When shopping for an item usually there is a warning like ‘order in 10 minutes and earn %10 discount’. In this way, they create the perception that you will lose the chance for that discount and need to act out immediately. It terms of student behaviors, it is usually seen that some students study just not to fail or not to lose their ranks between their peers. Another thing related to CD8 is that we do not want to experience undesired events so we do our best to avoid them. If there will be a negative consequence of an action, then to avoid it we do certain actions.

Usually as in the case of other CDs, CD8 works together with other drives. As Chou (2015) summarizes, CD8 is straightforwardly utilized by all sorts of

organizations and systems. CD 8 generates Black Hat results such as a high sense of urgency and obsession. Teachers, for example, threat their students with low marks and try to motivate them. In this way, students probably will do what is said to them not because they like it or want it but just they are scared and do not want to lose what they have. However, in the long-run, CD8 puts the user in a state of discomfort. And he finishes his article on his blog by stating that fear is an effective factor in motivating people. However, hope is what leads people to their goals.

Left Brain (Extrinsic) vs Right Brain (Intrinsic) CDs in Gamification. It would be necessary to state that Chou words the terms left and right brain not because their physical position in the brain but because they are easy to use and they symbolize logic and emotion. For him, the names given for this kind of distinction is nothing more than a semantic issue. However, he states that traditionally Left/Right Brain framework structure can be utilized to differentiate extrinsic and intrinsic motivation. In this respect, he makes the distinction as in the following: The Left-Brain CDs involve tendencies related to logic, ownership, and analytical thought. They are expressed in the following three CDs: CD 2, Development and Accomplishment; CD 4, Ownership and Possession; CD 6, Scarcity and Impatience. The Right Brain CDs are characterized by creativity, sociality, and curiosity and as illustrated by the following: CD 3, Empowerment of Creativity and Feedback; CD 5, Social Influence and Relatedness; CD 7, Unpredictability and Curiosity.

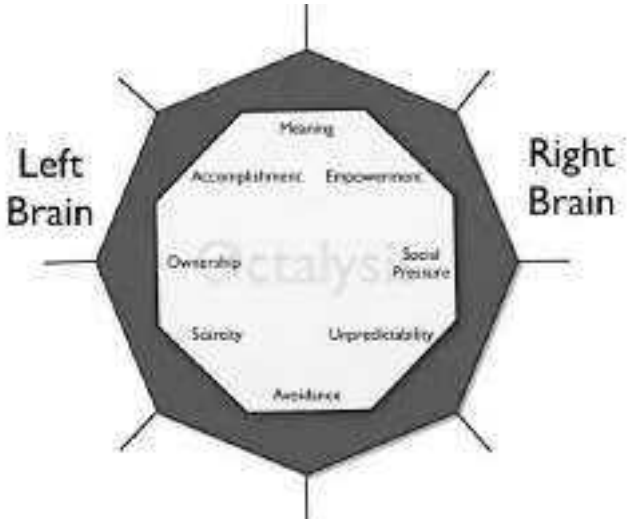


Figure 3. Left versus right CDs (Chou, 2015; p. 29).

As can be understood, left-brain CDs are related extrinsic motivation whereas right brain CDs are about intrinsic motivation. Chou (2015) points out that rewards, goals and purposes lead to extrinsic motivation. That is, although the task at hand is not so appealing and interesting, people may become motivated just because of the reward or the goal itself. For example, a daily core at work may be quite boring and demotivating but when offered extra payment, one can become motivated to get the reward which is an extrinsic. When extrinsic motivators are removed, the motivation can drop even below before.

On the other hand, when one enjoys doing an activity itself, it means that intrinsic motivation plays a role. A child's enjoying the painting can be a good example for this kind of motivation. In this respect, it can be said that left CDs are goal-oriented whereas right ones are experience-oriented and extrinsic motivation focuses on results whereas intrinsic motivation focuses on process.

Since there are disagreements on extrinsic and intrinsic motivation, Chou argues that when the reward or the objective is removed, if the person is still motivated that it is intrinsic motivation.

White Hat vs Black Hat. The White Hat CDs are represented by the CDs at the Top of the Octalysis diagram: CD 1, Epic Meaning and Calling; CD 2, Development and Accomplishment; CD 3, Empowerment of Creativity and Feedback. The Black Hat CDs are represented by the CDs at the Bottom of the Octalysis diagram: CD 6, Scarcity and Impatience; CD 7, Unpredictability and Curiosity; CD 8, Loss and Avoidance.

White hat and black hat CDs distinction is one of the original things that can be found in Octalysis design. The starting point for Chou (2015) is the difference between the modern games and the classical ones. He asks himself how the successful games that have caused addiction for months have ended up with losing many of its users whereas conventional games like chess or crossword puzzles continue to be popular for hundreds of years. As a result of his research and observations, Chou (2015) explains that through the final stages of the game experience there are significant differences between the way of classical games are designed and the way CDs motivate users.

Figure 4. White hat versus black hat CDs (Chou, 2015; p. 31).

One of the striking things that Chou realizes is that games that have gone viral within a huge population but have short lives utilizes the CDs that result in obsession, urgency and addictiveness. In this kind of games, people first glue to the games. However, towards the Endgame Phase they lose their enthusiasm and they just go on playing because of CD8. That is, they continue playing not because they enjoy themselves but because of the investments they have made so far. Eventually, some people begin to give up and due to Social Influence and Relatedness, other people give up playing. On the other hand, in the games that are timeless people still enjoy playing them and satisfied even in the Endgame Phase. Therefore, Chou labels some of the CDs as white hat and some other as black hat CDs. He also states that it would not be right to say that black hat drives are bad and white hat ones are good. When used appropriately, all of them can be beneficial.

When White Hat CDs are applied, users feel they get the power and satisfied with what they do in an autonomous way. In contrast, Black Hat CDs are usually related with obsession, anxiety and addition. When used, it is seen that black hat drives are very successful in motivating people. However, as time passes, people start to feel that they are losing the control and become demotivated.

When the above statements are taken into consideration, it can be thought that white hat drives are obviously more advantageous and will be enough if they are put into action alone. But they have a drawback which affects their efficiency. They cannot create a sense of urgency which can be created by employing black

hat drives. That is, without integrating black hat drives it would be hard to initiate actions only by white hat CDS (Chou, 2015).

The distinction between black hat and white hat CDs can help to explain or predict the success or failure at different stages. Why people start to do something with great excitement but eventually leave the design in subsequent stages can be explained in this respect. Chou (2015) explains that black hat means that when a system is designed well with black hat CDs, people can be motivated and act, and still be autonomous. But there should be black hat as triggers. According to Chou, the trick is not to force people beyond their skills. When black hat CDs are used too much and people are forced, the task will be carried out anyway but not in the long run.

The 4 Phases of a Player's Journey. So far, the eight CDs on the Octalysis and the distinctions made by using left and right brain dichotomy and black hat and white hat dichotomy have been investigated. Although these are very useful to explain a lot of things, they are not enough. Because the reason of doing something on the very first day will not be the same on the day hundred. So, Chou (2015) states about the levels in his design. Level 1 includes the eight CDs and left/right and white/black dichotomies. In level 2 there are four phases: discovery, onboarding, scaffolding, and endgame (p.39).

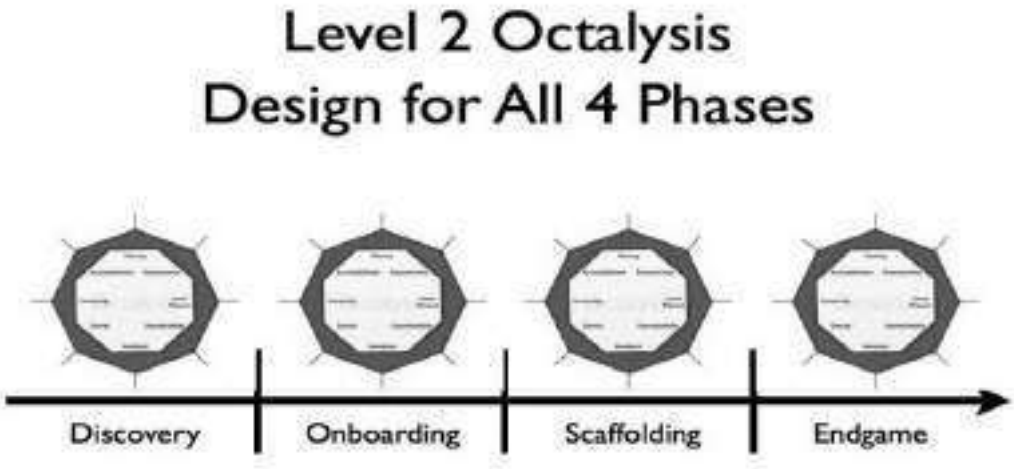


Figure 5. Level 2 Octalysis: Factoring in the four phases of a player's journey (Chou, 2015; p. 40).

It is possible to see situations in which a design attracts people at first but then fails to meet the expectations in the subsequent stages and causes people to drop out. Or on the contrary, a product may not even trigger users' motivation due to its too complicated design at the initial stage. Thus, people do not prefer to use. In this case, the system becomes demotivating and fails. So there needs to be a more human-focused design focused on four phases because as Chou (2015) points out user's interactions with a product continuously develops and changes throughout the process. Chou's four phases theory is an adapted version of Kenn Werbach's (2014) theory of identity, onboarding, scaffolding and mastery. But he modified these terms according to his own needs in Octalysis.

In phase one, which is discovery phase, a person starts to discover and learn for the first time as in the case of starting to learn a new language. When people decide to learn a totally new language which they do not know before, they are accepted as A1 learners (Common European Framework - CEFR). At this stage they try to get familiar with the language and try to learn basic vocabulary and sentences. However, when they continue studying the language, their needs start to change. They start to understand frequently used expressions and begin communicating in a simple way. That is, they pass to the second phase which is onboarding phase. According to Chou (2015) during onboarding, users become familiar with the rules, the options, the mechanics, and the win-states. When people become fully equipped and ready to go on, this phase ends.

In the third phase, which is scaffolding, the rules and things players learn in the previous phase become practical and they try to achieve more. If it is told with the language learning analogy, learners start to improve their language proficiencies and start to become more fluent. They become a part of B1 or B2 level learners. Scaffolding ends when a user believes that he/she has tried everything a product offers. They are now a veteran user ready for the end game.

Endgame, the fourth phase occurs when everything is done and there is nothing to find out anymore. In this phase, there is risk to lose motivation due to the difficulties in providing curiosity, meaning and excitement. However, once this stage is achieved, it is possible that people never gets bored.

People do not engage in a product not as a single event but through stages and develop experiences gradually. That is, a product may attract users in first place. Yet, it may become boring and uninteresting in time which results in failure in the design. Thus, each of the above-mentioned CDs needs to be evaluated for all these four phases.

Theories Underlying Gamification

Although gamification makes use of various scientific theories, what constitutes its foundations is behavioral economics, motivational psychology, self-determination theory, persuasive design and learning and development theories.

One of the most influential motivational theories that underlie gamification is Revee, Deci and Ryan's (2004) 'Self-Determination Theory'. A part of this theory is the concept of intrinsic and extrinsic motivation. The former refers to the type of motivation resulted from the inherent interesting or enjoyable nature of the task at hand. The latter occurs when individual does something because of outcomes.

There are studies on the role of Self-Determination Theory in gamification. In one of these studies, Muntean (2011) states that it is possible to use both intrinsic and extrinsic motivation together within gamification. In a gamification application, it is possible to employ extrinsic motivators such as levels, points and badges. On the other hand, it is also possible to use intrinsic motivation elements such as achievement, mastery, autonomy. Ryan, Rigby and Przybylski (2006) point out that it is possible to see the use of self-determination theory in video games. So, it will not be wrong to say that this theory can be used for gamification as well.

Since PBLs are the most frequently used game elements and they are seen as extrinsic motivators, there are a good number of studies which embrace Self-Determination Theory and investigate the use of PBLs in different contexts. For example, Cruz, Hanus and Fox (2015) carry out a study in which they investigate the impact of badges among video gamers. They argue that although badges are seen as extrinsic rewards that decrease players' intrinsic motivation according to the self-determination theory, they indicate that these may be seen as intrinsic motivators by some players. In other words, they find out that badges can improve the motivation of interested players and their enjoyment, engagement and time they spent.

There are several other studies in which researchers have reached conflicting findings in terms of the use of PBLs. Mekler, et al. (2017) empirically investigate the effect of gamification on intrinsic motivation and need satisfaction. They find out that game elements such as points, levels and leaderboards do not significantly affect competence or intrinsic motivation. But they do not impair intrinsic motivation contrary to the previous research findings (Hanus & Fox, 2015). They suggest that extrinsic motivators like points, levels and leaderboards can promote performance quantity. Gåsland (2011) finds out that a gamified system using points-based game elements is perceived to be somewhat motivating and engaging by students. But there are also some studies that do not mention positive effects of using PBLs. Meyer (2008) states that the majority of the students who take part in the study find the use of points ineffective in term of the quality of their postings.

In their study in which they investigated the use of points as game elements, Attali and Arieli-Attali (2015) find out that points do not have any effect on students' achievement levels but on their speed of response.

Another well-known motivation theory that underlie gamification is Csiskzentmihalyi's (1990) 'Flow' theory. According to Csikszentmihalyi (1990) when people are in flow they are absorbed by the task at hand and lose track of time. They isolate themselves and experience a heightened motivation. Like self-determination theory, flow theory is also applied to video games successfully (Hsu & Lu, 2004; Johnson & Wiles, 2003). In addition, researchers such as Chen, Wigand and Nilan (2000) and Pilke (2004) argue that while interacting with technology, it is important to design flow experience. In this vein, it is possible to see gamification studies which focus on the Flow experiences. For example, Li et al. (2012) and Dominguez et al., 2013 use "flow theory" as the foundational theory of gamification in their studies.

Dominguez et al. (2013) argue that 'flow experiences' cannot be applied to traditional educative content without entering the field of edutainment or serious games. In their study which is aimed to reveal the correlations between flow experiences and cognitive loads in game-based and non-game-based learning groups, Chang, Liang, Chou and Lin (2017) find out that when multimedia learning materials are designed by taking games interests and interactions into the consideration, then learners may not only live more flow experiences but enjoyment.

However, Self-Determination and the Flow theories are not the only theories mentioned in gamification studies. There are other researchers who try to expand theoretical foundation of gamification. According to Hamari (2013), some of the several theories underlying gamification are the goal setting theory (Locke & Latham 1990), social comparison and influence, social proof theory and social validation. Similarly, Landers, Bauer and Callan (2017) state that goal setting theory plays a significant role in understanding the success of certain game elements such as leaderboards and that to better understand the effect of gamification, the goal setting theory and other psychological underlying theories should be investigated. As Bandura (1997) points out, this theory assumes that setting goals can increase performance in three ways. Setting goals help people increase their expectations and their motivation can increase. In addition, when goals are assigned, this may increase self-efficacy. Finally, when people achieve their goals, they experience high levels of satisfaction and their future performances for the same activities increase. Goals being context-related, immediate feedback to users can strengthen these effects. Clear goals are at the center of the flow theory (Csikszentmihalyi, 1990). So, it would not be wrong to assume that one of the several underlying theories is the flow theory. If goals are clearly set and immediate feedback are given to the users, the change of getting into the flow increases. Of course, this is true if the skill-challenge balance is at optimum. According to Locke and Latham (1990) only clear goals may not be enough to reach the desired goal, users should also commit themselves to achieve those goals, which is called as goal commitment.

Another rationale behind gamification is social comparison. That is, people tend to compare themselves to the other in the same social circle and benchmark themselves accordingly. Hamari (2013) find out that when users are socially influenced and recognized through gamification, this increases the popularity of gamification applications. On the other hand, social proof theory (Cialdini, 2001) predicts that when people recognize that other people start to engage in behaviors, then they are inclined to engage in those behaviors more. In addition, people evaluate others' preferences and try to make valid choices according to their social environment to show their conformity to the behaviors and expectations of others.

Finally, recent studies have shown that it is possible to use Directed Motivational Current Theory (Dörnyei et al., 2016) instead of Flow Theory

(Csikszentmihalyi, 1990) because as Dörnyei et al. (2016) argue, Flow Theory has problems when applied in longer periods and may fail to explain the experience when there is more than one activity to be completed. Since language learning is not confined to one single activity to be completed in a single time period, it may be possible to integrate DMC into the Octalysis to get more accurate results.

In the following section, the most important theories underlying gamification are going to be investigated in a more detailed way.

Self-determination theory (SDT). Deci and Ryan (2012) criticize historical and contemporary motivation theories. They state that these theories mostly accept motivation as a unitary concept and they do not clearly put forward what types of motivation exist. Even when they clearly state the types of motivation, they try to reach a total motivation type as Bandura (1997) did. Deci and Ryan (2012) argue that these theories are successful in predicting the amount of behavior but they ineffectively predict the quality of behavior. They state that what is important is to be able to determine whether one individual is motivated in an autonomous and controlled way rather that emphasizing the quality of engagement. And they develop the theory of Self-Determination.

According to this theory, there are three core psychological needs to be met in SDT. These are competence which means that one has the mastery and can influence the outcome; autonomy which means one is guided form inner drives rather than outside ones; and finally, relatedness which means being connected to others (Ryan & Deci, 2000). According to Deci and Ryan (2012), satisfaction of these facilitates autonomous motivation. And prevents individuals from feeling the need to get motivated in a controlled way or totally losing the motivation which is also called as amotivation or lack of intentionality.

In terms of autonomous and controlled motivation, SDT posits two assumptions. Accordingly, people are motivated doing an action because it is interesting and enjoyable. And doing the activity itself is the reward like a child playing with toys. This is called as intrinsic motivation. If this is interpreted in educational context, as Grolnick and Ryan (1989) and Benware and Deci (1984) argue, it can be said that when intrinsic motivation is dominant, students' learning tends to become deeper and conceptual. Also, they have longer memory.

On the other hand, when a separable consequence is the reason for doing an activity, it means extrinsic motivation exists. Among classical extrinsic motivation, there are rewards and punishments. When people's behavior is controlled with rewards or threats, their behavior tends to become dependent on the contingencies. That is, as Peterson (2006) points out, instead of paying attention to the external rewards or punishment, it is important to undertake activities because of their own appeal.

In this respect, since, as Csikszentmihalyi (2008) puts forward, Flow is a subjective combination of intrinsic motivation and confidence that comes from complete immersion in a task, it can be said that intrinsically motivated behaviors are also important in terms of creating opportunities for Flow experiences. Ushioda (2008) states that when people intrinsically motivated, they concentrate to learn for its own sake. This means that there is an optimal skill-challenge balance which results in higher levels of involvement in learning activities, engagement in more efficient and creative thinking process. They can also employ more problem-solving strategies.

According to Deci and Ryan (2012), autonomy can be defined as the capacity for and desire to experience self-regulation and integrity. Being a healthy person is correlated with being autonomous. That is, the greater the autonomy one develops, the healthier personality can become possible. In addition, autonomy includes internalizing and integrating external regulations over behavior. When a person develops autonomy, it becomes easier to control the drives and emotions. In this way, a person can maintain the intrinsic motivation.

In time researchers have developed various sub-theories to cultivate SDT. One of these is Organismic Integration Theory (Ryan & Deci, 2000). In this theory, Ryan and Deci (2000) describe extrinsic and intrinsic motivation on a continuum. They state that depending on the level of regulations there are multiple types of extrinsic motivation: external regulation, introjected regulation, identified regulation and integrated regulation.

When behaviors are externally motivated to satisfy the external rewards or demands, this is called external regulation. When individual's perceived locus of causality is not completely but somewhat from external sources and individual does

the behavior to satisfy his/her ego, this is called introjected regulation. When the individual's locus of control is somewhat internal and individual sees the behavior as important and perceives that he/she is doing the behavior because of his/her internal forces then this is called as identified regulation. Finally, when the reason to do the behavior stems completely from internal forces and is associated with other values, it is called internal regulation. But it should be pointed out that internal regulation is still a type of external motivation no matter how it is internally forced. Because still the individual does not do the action only to satisfy his/her psychological needs.

In another theory which is called as Cognitive evaluation theory, Ryan and Deci (2000) focus on the competence and autonomy during an activity. Accordingly, individual's intrinsic motivation is affected by his/her perception of rewards or feedbacks. When the individual perceives the rewards or feedbacks as controlling, then his/her intrinsic motivation decreases. However, when he/she perceives them as informational, then this can enhance their intrinsic motivation because he/she believes that his/her behavior is driven by the self. In this respect, as Cruz et al. (2015) argues, whether badge systems are seen as controlling or informational is unknown and this may influence individual's intrinsic motivation.

In addition, another theory, Signaling Theory, attempts to explain how communicators are sent and interpreted particularly transmitted cues (Donath, 2007). That is, badges or other game elements as communicators may have various purposes except form their initial goal which is motivating the individual. They may not be only meaningful for the individual who possesses but may also can give information about the individual to the others. Other people can interpret these for assessment, strategic or conventional purposes.

Flow Theory. In literature, the Flow experience is interpreted within the scope of Positive Psychology which aims a change in psychology. Seligman (2002) argues that there is a need for a change in psychology. Studies have focused on the negativity and the problems. However, what is important is to focus on the best qualities in life rather that the worst things. International Positive Psychology Association (IPPA) which was founded in 2007 defines positive psychology as "the scientific study of what enables individuals and communities to thrive".

There are three research areas in positive psychology: positive subjective experience, positive individual traits, and positive institutions (Seligman and Csikszentmihalyi, 2000). The authors state that “Flow”, which is initially called as “autotelic” experience by Ciskszentmihalyi, is a key construct area in positive subjective experience. Csikszentmihalyi (1990) points out that the moments in which one get into a special state of absorption and enjoyment is called as flow experience. In other words, when one is totally involved in a task at hand, it is possible to mention the existence of flow experience. In this respect, it is possible to link flow with intrinsic motivation.

According to Csikszentmihalyi (1990), flow can occur at two levels: micro and macro. At micro level, there are experiences that occur in everyday life. On the other hand, at macro level, experiences associated with higher levels of complexity and demand exist. It can be said that these kinds of sophisticated experiences are very similar to the Maslow’s (1968) “Peak Experiences”. Maslow identifies 14 characteristics of these experiences. Among these are feeling of being detached from concerns, strong concentration, egoless and unselfish perception, disorientation in time and space, feeling of being meaningful, beautiful and desirable. In this respect, as Jackson (2000) states, flow and peak experiences are similar. However, they are distinct concepts. Flow model incorporates both affective and cognitive component and grounded in multidimensional theory of optimal experience.

It can be said that flow is internal and conscious process and includes several simultaneous positive experiences (Jackson & Csikszentmihalyi, 1999). In flow experiences, one totally focuses on the task at hand and is absorbed by the activity. While in flow, a person ignores all other thoughts and emotions and mind and body work effortlessly. In this respect, Csikszentmihalyi (1990) and Nakamura and Csikszentmihalyi (2009) point out that there are nine dimensions of flow which are challenge-skill balance, unambiguous feedback, action-awareness merging, concentration on a task, sense of control, loss of self-consciousness, time transformation, autotelic experience. These nine dimensions represent optimal psychological state of flow.

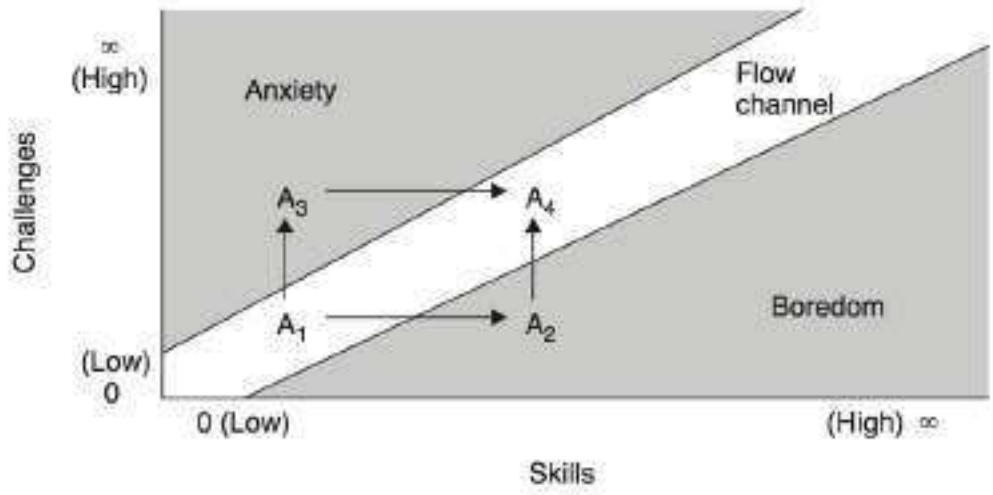


Figure 6. The Flow (Csikszentmihalyi, 1990; p. 74).

As it is seen in the Figure 6, the level of flow can be determined by the challenge and skill balance (Csikszentmihalyi, 1990; p. 74). When a task is too difficult above individual's skill level, it causes anxiety. As the difficulty of the task is tuned according to the skill level of the individual then it is possible for that individual to experience flow. However, as the task becomes easy and becomes less challenging, then individual starts to become demotivated and boredom occurs.

There are countless studies devoted to the investigation of flow experiences in many areas. In educational context, Karageorghis, Vlachopoulos, and Terry (2000) state that flow experiences among school students can be facilitated when well defined personal goals which are attainable but challenging are set, students are given choices and skill-learning techniques are used.

When the Flow Theory is interpreted with Octalysis lenses, it is seen that what is aimed with Octalysis is to enable individuals to experience the flow. All the CDs Chou (2015) mentions are used to optimize the circumstances in which individuals can get into flow.

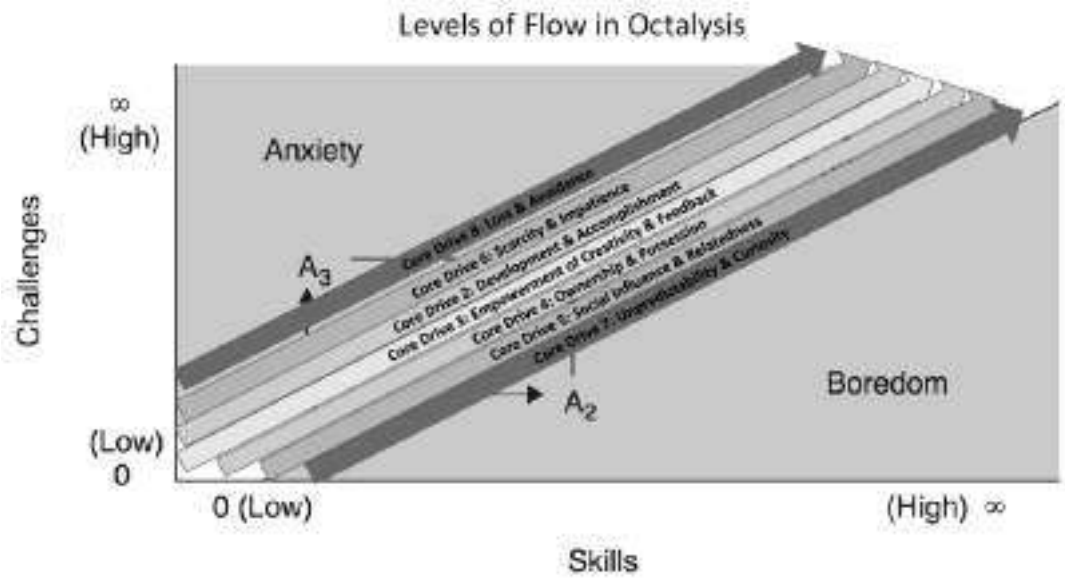


Figure 7. The Flow from Octalysis perspective (Chou, 2015; p. 418).

When the above skill-challenge graph proposed by Csikszentmihalyi (1990) is reinterpreted with Octalysis (Chou, 2015; p. 418), it is seen that CD 8: Loss and Avoidance is placed at the top of the Flow. Because when people struggle not to fail and lose something, it means that the task is very difficult and above their skill level and this results in increased anxiety. In a classroom environment, when the difficulty level of the lesson is above students' skill levels, students only study not to fail in the exam. When the skill level of the task is slightly decreased, individual enter the zone of CD 6: Scarcity and Impatience. Here, the difficulty level is still above individual's skills level and as Chou (2015) states, users still feel challenged and occasionally frustrated. However, if they manage to act immediately, there is a chance to overcome the challenge.

Below the CD 6, it is seen that there is CD 2: Development and Accomplishment. At this zone user feels a moderate challenge and feels accomplished and component. This is where users start to experience the Flow and achieve their goals. In this way, they also start to improve their confidence. In classroom, it is possible to mention about a balance between students' skills and the challenge of the lesson. In this zone, students know that they can achieve and thus, they can enter the Flow and start to study.

At the center lies CD 3: Empowerment of Creativity and Feedback. According to Chou (2015) most people enter Flow when CD 3 is accomplished. Users'

creativity levels are at most and they adjust their strategies to discover better ways to do things. Especially when there is a rotation between CD 2 and CD 3, there will be better chances to experience Flow.

On the other hand, below the CD 3, there is the zone of CD 4: Ownership and Possession where the difficulty levels start to decrease and users feel relaxed in mind since there is little challenge. People start to collect, organize and put things together in a mindless way. In terms of students, excitement vanishes and engagement decreases. At this stage in order not to cause boredom, proper designs are needed. Below CD 4, there comes CD 5: Social Influence and Relatedness. At this zone, people do not need to have too much skill to complete the actions. The activity can still be fun, but there is almost no challenge. But, when they feel relatedness, acceptance and bonding with others, people still find the activity pleasant and feel relaxed.

At the bottom, there is the CD 7: Unpredictability and Curiosity. This drive adds chance and randomness to the experiences and this makes the activity easier. Users do not need to have skills to do the action. They can take the desired action and just wait for the results.

As can be seen, in Flow there is no place for CD 1: Epic Winning and Meaning. Chou (2015) argues that CD 1 is missing because it is “completely out of the context of player skill level compared to challenge difficulty” (p. 412). Therefore, he states that CD 1 does not really fit into Flow model but may constitute the reason why people engage with the experience. Similarly, Dörnyei et al. (2016) argue that there is no place for CD 1 or meaning in Flow model. Thus, they propose DMC instead of Flow.

Directed Motivational Currents (DMCs). In any context from time to time, it is possible to find individuals in a kind of heightened motivational state which results in intensive state of focused productivity while working on something. To explain this, Csikszentmihalyi (1975) introduces the concept of Flow experience. Csikszentmihalyi (2008) further explains that when people are in the Flow which is a subjective state they forget the track of time, do not realize that they get tired and isolate themselves from everything except the activity. In this respect, it is possible to identify Flow experiences in many fields. However, in language learning context,

Dörnyei et al., (2016) criticize Flow Theory and propose that DMCs can replace Flow. According to Dörnyei et al. (2016), a DMC “is an intense motivational drive – or surge – which is capable of stimulating and supporting long-term behavior (such as learning of an L2)” (p.2). However, they warn that a DMC is not equivalent to high levels of motivation in general. It is relatively short-term, highly intense and goal oriented. Besides, it is not the same as the Flow experience

Flow Versus DMC. Dörnyei et al. (2016) do not redefine what Flow experiences mean. Instead, they interpret flow in a similar way to Csikszentmihalyi (1975). Dörnyei et al. (2016) acknowledge that Flow is a subjective state of intense involvement in a task. In a flow, people feel so absorbed into the activity that they compare it being outside of everyday reality. One of the most distinguishing feature of Flow experience is that people focus on an activity in a concentrated way and enjoy doing the activity in which they forget about the time and other elements of human experience. Dörnyei et al. (2016) also add that Flow experience can be observed in both artistic and non-artistic environments such as playing digital games.

When the Flow and DMCs are compared, it is seen that there are some similarities between these two concepts. According to Dörnyei et al. (2016), one of the similarities is that total absorption characterizes both Flow and DMC. However, Flow and DMC explain this total absorption differently. Whereas Flow most likely occurs in structured environments (Csikszentmihalyi, 1988), DMC involves longer periods of total absorption. Another difference is that Flow concerns shorter time window when compared to a DMC and mostly it is limited to a single activity. Yet, a DMC can take place for longer time periods and includes multiple related tasks.

In terms of sustaining the flow experience, Ciskszentmihalyi (1990) states that “one must increase the complexity of the activity by developing new skills and taking on new challenges” (p. 30). That is, as Dörnyei et al. (2016) point out, new challenges should be continuously added to the activity to lengthen the time span of flow experience. Usually, this is applied to digital games in which the length of the game is extended by adding new levels and stages. Dörnyei et al. (2016) criticize this by stating that even the time span of the Flow experience is extended with new levels and challenges, this does not change the reality that the person is still in the same activity.

In this respect, DMCs differ from Flow experiences. Dörnyei et al. (2016) state that “in DMC diverse tasks and task components are subsumed by an overarching current, linked together by the ‘directedness’ of the process as the current moves toward an overall goal or vision” (p. 5). In other words, in Flow experience at the heart of which lies autotelic experience there is an end or purpose and moving toward a larger goal is not emphasized. Because of the intrinsic motivation, individual enjoys doing the activity itself. However, approaching a desired long-term target or vision constitutes the DMC concept (Dörnyei et al., 2016). This means that not all the tasks should be necessarily enjoyable. They are rewarding because they allow the individual to move toward a highly valued end goal.

To sum up, in language learning processes, which can last for years, gamification applications based on Flow Theory (Csikszentmihalyi, 1975) may cause some problems in terms of explaining the long-term nature of language learning and striving for an ideal goal. In addition, it may not always be possible to provide multiple tasks which provide autotelic experiences for students in language classrooms. Throughout their language learning experiences, students will inevitably come up with tasks which are not perceived as enjoyable. However, when holistically taken into consideration, all tasks provide enjoyable experiences together while they transfer individuals to a highly valued end goal. Thus, in the following section DMCs will be investigated in detail.

***Development of D*MCs.** L2 motivation is a vast area and there are many studies underlying the Directed Motivational Currents Theory. Dörnyei et al. (2016) suggest that to better understand the concept of DMCs, the historical development of motivation and L2 motivation research that leads to the development of DMCs should be investigated.

It is known that L2 motivation research has followed the mainstream motivation research. Williams and Burden (1997) state that early works on motivation research are in line with the thoughts of behaviorists. As behaviorists, they try to explain the reason why people behave by investigating the animals in laboratory environments. They conclude that human motivation to learn is closely knitted to the rewards and reinforcements provided and the degree to what biological needs are met during the early years. They found out that rewards systems are the most effective tools to motivate people. Which means that from a

behaviorist perspective motivation is largely considered in terms of external rewards. Based on behavioristic theories of motivation, drive reduction theories dominate the motivation research and theories. As a kind of reconceptualization of drive theories, Atkinson's (1964) achievement motivation seems promising for some time. However, soon it is realized that knowing individual's level of need to achieve is not enough. It is also important to know the strength of tendency to avoid the task.

But these early views of motivation are found to be too simple to explain the human behavior. In addition, lack of consciousness is another problem that these early theories fail to explain. This view which places value to the conscious control paves the way to the development of cognitive psychology. One of the most important issues in cognitive view of motivation, is the role of choice. That is, people have choice about how to behave and thus could control their actions. To make choices, people need to set goals for themselves and strive to achieve these certain goals. Therefore, as Williams and Burden (1997) state, from a cognitive point of view, motivation is related to why people decide to act and what factors affect their choices. But, cognitive perspective has its own drawbacks. One of the most important problems that cognitive psychology bears is its failure in explaining the roles of affective factors, emotions and social and contextual influences.

This gives rise to the development of first constructivist approach and then socio-constructivist perspective. Constructivist perspective puts emphasis on the reality that each individual is motivated differently. Because people make their own sense of the surrounding world different from others and they act on their internal disposition (Williams & Burden, 1997). Their reasons to learn is different from each other. In addition to constructivist perspective, researchers have found out that a person's motivation is also closely related to social and contextual influences. The culture, context and the social situation can all affect their motivation level.

When L2 motivational research is investigated it is seen that the historical development of it follows the above-mentioned perspectives. A brief historical development of L2 motivation research is well documented in Dörnyei et al.'s (2016) book. According to Dörnyei et al. (2016) it is possible to divide historical development of L2 motivation research into four main phases. In the first phase which is called as 'social-psychological period' by Dörnyei and Ushioda (2011), motivation was investigated with a macro point of view. In this sense, individual

learners mostly were not at the center of focus. Throughout this period studies by Gardner and his friends played leading role in effecting the motivational studies (Gardner, 1985; Gardner & McIntyre, 1993). One of the most salient dichotomies investigated in this period was Integrative versus instrumental orientation or motivation. One of most notable outcomes of the researches carried out in this period is that the concept of L2 motivation was acknowledged as an important ingredient of successful SLA.

The next period which is called as 'cognitive situated period' witnessed many studies with a new perspective. In this period, studies by Crookes and Schmidt, 1991; Dörnyei (1994), Julkunen (1989); Oxford and Shearin (1994); Skehan (1989); and Ushioda (1994) came to the fore. Dörnyei et al., (2016) point out that in previous period the motivational studies were carried out by psychologists who were interested in SLA. However, in this period, it can be said that, SLA researchers who are interested in psychology started studying L2 motivation. In addition, it can also be argued that the fundamental point of view held by the psychologists who studied L2 motivation changed in this period and researchers started to adopt more micro perspective. In this way, they contributed the foundation of the field with several theories they developed. They transferred cognitive theories from educational psychology such as self-determination theory, goal theories and attribution theory. In this respect, what it can be stated that they tried to unveil what underlies the instructed SLA.

Having realized that L2 motivation was not so simple and consisted of many interrelated and complex factors, researchers started to investigate the learner, the learning task and the learning environment in socio-dynamic period. In this period, Dörnyei's (2005) thoughts about L2 motivation opened new areas. At the centre of Dörnyei's approach was learner's self-concept. Dörnyei (2010) stated that what Gardner's 'integrative motive', which has its roots in possible selves theories (Markus & Nurius, 1986) and self-discrepancy theory (Higgins, 1987), can be developed in a way which included possible self and he proposed the 'Ideal L2 Self' concept. Soon this new concept got many researchers' attentions and learners' future self-images began to be investigated.

Then, these future self-images turned into the concept of 'vision' which means vivid mental image of the experience of successfully accomplishing a future

goal. Later, this concept extended by Dörnyei et al., (2014) and the concept of DMC which is a vision related phenomenon was founded.

***Underlying theories of DMC*.** Since traditional motivational theories have failed to explain the relationship between motivation and the subsequent behavior and the temporal characteristics of motivation, a new construct of motivation is needed (Dörnyei et al., 2016). Traditionally, theories that explain motivation separate motivation and the behavior that occurs afterwards. Another issue that traditional theories were unable to answer is the temporal side of motivation. It is known that language learning takes years and motivation that leads people to learn is not steady instead it has ups and downs throughout the learning experience.

Dörnyei et al. (2016) mention three theories that have relevance with DMCs in terms of time. These are time perspective, contingent path theory and velocity in goal pursuit. Time perspective theory (Zimbardo & Boyd, 1999 and de Volder & Lens, 1982) focuses on people's understanding of time by focusing on the past, on the present and on the future. According to this theory, future oriented students show better performance in attaining their academic goals (Kauffman & Husman, 2004). When people can imagine the results of their current behaviors in the future, they engage in activities more willingly. With this theory and its findings, Dörnyei et al. (2016) could link present and future. Contingent path theory (Raynor, 1974; and Raynor & Entin, 1983) assumes that achievement of a task depends on the achievement of preceding tasks. Completing tasks one by one to reach the final goal means that there is a contingent path which is made of several steps. Finally, the velocity theory explains the rate at which people approach their goals. According to Dörnyei et al. (2016) the concept of velocity represents the dynamic nature of motivation

Dörnyei et al. (2016) also state that approach and avoidance dichotomy in motivation is directly related to the understanding of DMCs. In traditional understanding, a motive results in behavior. In this respect, it can be said that a behavior can be motivated either by external or internal resources. In terms of approach and avoidance motivation research, Dörnyei et al. (2016) state that internal resources that cause behavior are directly related to the vision they mention in DMCs construct. In motivational psychology, the distinction between external and internal forces that cause motivation is made in approach versus avoidance

motivation (Elliot, 2008). When this dichotomy is taken into consideration, approaching a desired target is at the heart of DMC. Dörnyei et al. (2016), who define DMCs “as the optimal form of approach motivation energizing sustained behavior” (p.34), state that developed forms of DMCs can partially be seen as long-term approach motivation.

Another issue that is directly related to DMCs is the notion of resilience. Dörnyei et al. (2016) point out that long term behavior energized by DMC “…involves regulation that does not rely on disciplined scaffolding by a strong will or personal resilience” (p. 34). This is related to the positive side of DMCs and resembles flow in this sense. Jackson (2012) states that fear of failure does not exist in flow and people in flow positively approach the challenges.

Another point Dörnyei et al., (2016) pointed out in their book is that a DMC is always directional. That is, there is always a specific end-goal in DMC and this directedness is known throughout the DMC experience. In this respect, they state that this is one of the differences from the flow theory. Since DMC is an optimal form of approach motivation, this directed aspect of DMCs cannot be a surprising issue because for an approach to happen, the end-target should clearly be defined so that the goal can be approached. In other words, without setting the goals and they are “pursued volitionally” (p. 41), action cannot take place. In other words, another underlying theory of DMC is the Goal Setting Theory (Locke & Latham, 1990).

Goals have long attracted the interest of researchers from motivational psychology. Dörnyei et al. (2016) summarize five main findings of decades of research which was inspired by Locke and Latham’s (1990) “goal setting” theory. Accordingly, researchers have found out that the more difficult the goal, the greater the achievement. In addition, the more specific or explicit the goal, the more precisely it can regulate performance because goals that are both specific and difficult lead to the highest performance. Moreover, commitment to goals is most critical when goals are specific and difficult (i.e., when goals are easy or vague it is easy to inspire commitment; it does not require much dedication to reach easy goals, and vague goals can be easily redefined to accommodate low performance). High commitment to goals is attained when (a) the individual is convinced that the goal is important; and (b) the individual is convinced that the goal is attainable (or that, at least, enough progress can be made toward it) (p.41).

According to Dörnyei et al. (2016) the goal setting theory lacks in answering several questions. For example, what makes the goals important cannot be answered. So, in order to be able to better reflect the intensely and purposefully directed aspect of DMCs, they suggest three key extensions to goal setting theory which are vision, self-concordant goals and proximal sub-goals.

Vision. To explain long-term motivation researchers have tried to conceptualize the higher order factor that results in sustained motivation and the concept of vision has emerged within L2 motivation research. For the first time, the notion of vision is used by Markus and Nurius (1986) in possible-selves theory. They did not explicitly used the term, though. The theory of possible-selves helps us understand something about our projected future self. What kind of person we might and would become and what kind of person we are afraid of becoming are investigated in the theory. According to Higgins (1987), there is a gap between our understanding of who we are now and who we would like to become in the future and people positively take action and try the close the gap. By bringing the findings of motivation research together, Dörnyei (2005, 2009) developed L2 motivational self-system theory. This theory helps us understand the way people imagine themselves in the future is important since this vision will energize their present behaviour. There are three key components of this theory: The ideal L2 self, the ought to L2 self and L2 learning experience. The role of imaginary and visualization aspects of L2 motivation has recently been investigated by some researchers such as Dörnyei and Chan, (2013), Kim (2009), You and Chan (2015). It is also possible to find studies on imaginary intervention (e.g. Chan, 2014; Macay, 2014; Magid, 2014; Sampson, 2012). In short, Dörnyei et al. (2016) state that

> "the concept of vision is formed as an extension of an abstract, cognitive goal, on which strong sensory elements – that is, tangible images related to actually achieving the goal – are super imposed. This works not only in connection with the final goal, but is also interconnected with each individual sub-goal along a DMC patway" (p.47).

Self-concordant goals. As Markus and Ruvolo (1989) argue, a vision can be seen as a highly-personalized goal. However, according to Dörnyei et al. (2016), to say that a person really has that goal, it should be determined by the individual himself/herself and needs to represent the individual's interests, passions, values and beliefs. Sheldon and Elliot (1999) call this kind of goals as self-concordant goals

which are at the center of DMC understanding. Because when a goal is aligned with someone's personal identity, it initiates a motivational current by making the goal more attractive for that person. So, it can be said that in DMC during the pursuit of self-concordant goals, people willingly enter into the action and experience the feeling of well-being. In addition, the sustained energy is continually regenerated over time.

Primary conditions of DMC. How does a person get into a DMC? According to Dörnyei et al. (2016), DMC is triggered by a specific event or by a group of individual factors. First of all, for a DMC to occur, necessary conditions need to be aligned and there should be a trigger. These are a clear set of goals, sense of ownership and control, perceived balance of challenges and skills, openness to the DMC experience.

For DMC to occur, goals need to be clearly defined and set because they strengthen the structure of DMC. When goals are set clearly, this means that persons will behave consistently to attain that goal. In addition, if a person does not feel that she/he has the control over the process and owns the outcome, then she/he cannot get into DMC. Ajzen (1991) uses the term perceived behavioral control to explain this concept. Furthermore, there should be a challenge-skill balance for successful DMCs. Hefferon and Boniwell (2011) explain that if an activity is hard and above our skill level, we become anxious. On the other hand, if it is too easy, that is, below our skill level and produce no challenge then we get bored. For a DMC to occur, there should be this challenge-skill balance. Finally, according to Csikszentmihalyi (1975), some people experience flow more easily because they have autotelic personality. Nakamura and Csikszentmihalyi (2002) state that these kinds of people have certain metaskills and competencies which help them stay in flow. When right conditions and opportunities are provided, many people may experience a DMC.

It is known that it is possible to achieve big and meaningful things within DMC. Apart from this, it is also possible to live satisfying and enjoyable experiences which cannot be found in daily life. In other words, as Dörnyei et al. (2016) state, DMC results in positive emotional well-being and individuals caught in the current can enjoy themselves and feel well because of getting closer to their ultimate goal and seeing their progresses. This feeling of joy and enjoyment is not like those felt in

daily activities it “stems from a complete feeling of connectedness with an individual’s core understanding of who they really are” (p. 103). This well-being is among the core characteristics of DMC. Moreover, Selçuk and Erten’s (2017) study indicates that all core dimensions of DMC construct can be confirmed.

In this sense, according to Dörnyei et al. (2016), possible-self theories in social psychology, goal theories in motivational psychology and flow theory in positive psychology contribute to the understanding of positive emotionality of a DMC. In addition, another concept from positive psychology which is “eudaimonic well-being” should be integrated into these theories to better understand the unique quality of DMC. Dörnyei et al. (2016) state that especially eudaimonic experiences (Waterman, 2008) and dedicated effort in the pursuit of excellence and authenticity (Huta & Waterman, 2013) should be deeply investigated in this process.

Eudaimonic well-being (Huta & Ryan, 2010; Huta & Waterman, 2013; Ryan & Deci, 2000; Waterman, 2008) refers to a deep and often enduring sense of personal contentment, as contrasted with a more transitory euphoric and ‘in-the-moment’ experience of happiness. Eudaimonia is linked to the experience of one’s potential and the realization of personal fulfilment and in much of the research it is conceptualized at trait level. However, in DMC it is conceptualized at state level because DMC is interested in the positive emotions experienced in the context of engagement in goal-directed activities. Waterman’s (1993) eudaimonic identity theory investigates the state eudaimonia. According to this theory, activities that can generate feelings of eudaimonic well-being are those which involve the development of one’s skills and talents, the advancement of one’s purpose in living or both” (Waterman, 1993; p. 679). There are three core elements in this theory: purpose and meaning in life, dedicated effort in the pursuit of excellence and authenticity.

DMC path. In DMCs, a person knows when it starts and is aware of the final vision which means that there is a clear start and end-point within a DMC. Between these two there are three key elements that make up the path: “automized routines, subgoals and process checks and affirmative feedback” (Dörnyei et al., 2016; p. 81). For a DMC to occur, there should be a set of fixed routines that allow the emergence of motivational autopilots because as Aarts and Custers (2012) state, when the actions are repeated again and again in the same context and adjusted according to the goals, then they become habitual and there is no need for conscious control

of these goal-directed actions to achieve the goal. Another vital issue is the existence of subgoals. Thanks to subgoals, motivational energy can be transferred and they can also be used as progress checks in DMC because usually in language learning the end goal is in far away and subgoals and midrange goals can help individual to go on within the current to attain the ultimate goal.

In terms of affirmative feedback, Hattie and Timperley define feedback as "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (p. 81). Dörnyei, et al. (2016) state that it can be used for motivational and instructional purposes. Feedbacks can highlight the discrepancy between of the main attractor state and the desired goals. In this respect, Voerman, Meijer, Korthagen and Simons (2012) point out that there are two kinds of feedback: progress feedback and discrepancy feedback. The former which is also called as affirmative feedback focuses on the discrepancy between the initial and current state of performance. It gives information about what has achieved so far. The latter focuses on the current and the desired level of performance and gives information about what is needed to be achieved. In DMCs, mostly progress feedback plays role because as Dörnyei et al. (2016) stated "this forms a powerful type of progress check which makes progress toward the target feel real and achievable and which thus fuels subsequent efforts" (p. 93).

Group DMCs. Having laid out the theoretical foundations of individual DMCs, Dörnyei et al. (2016) argue about the possibility of group DMCs. What about the group level DMCs? Dörnyei et al. (2016) state that it is not so easy to explain DMCs at group level since as Dörnyei and Ushioda (2011) point out, some preconditions need to be met to investigate motivation in classroom level. First of all, teachers should behave appropriately and establish good relationships with their students. Then, there should be a supportive and pleasant environment in classroom environment. Finally, there should be a cohesive learner group with appropriate group norms. Also, it shouldn't be forgotten that these are not independent factors determining group motivation. In fact, they are related to each other and group dynamics as well. In other words, it is a complex and dynamic issue with many variables. However, it can still be said that DMC-related interventions can be used to create "optimal learning environments" and when these optimal learning

environments area enabled, motivational currents can be formed more easily in language classrooms.

Dörnyei et al. (2016) argue that although the idea of applying DMCs at group level seems difficult, it is not impossible. The question to be answered here is whether a group of students can get into a collective motivational hyperdrive. They add that once it is proven that DMCs can also function at group level, it can be possible to design interventions for DMC experiences tooccur in L2 environments. In this respect, it can be concluded that by the development of focused interventions, raising the collective motivation of the whole group can become possible.

In literature, there are some studies that focus on the heightened group motivations. For example, Barsade (2002) investigates the emergence of group level energy and states “processes of contagion play an important role” (as cited in Dörnyei et al., 2016; p. 142). When there is an emotional contagion, it effects the whole group and by creating positive emotional loadings and enthusiasm with a task, it is possible to trigger the positive mood within a group of people and this positivity effects the level of motivation of that group. Aarts and Custers (2012) talk about goal contagion. This means that when people understand the goals from the behaviors of other people, then their goal pursuit is triggered automatically. In short, it can be stated that when there is a goal that is worth, then whole group can be influenced (Papies & Aarts, 2010).

In addition to above mentioned studies, there are also other studies which investigate the use of flow theory at group level. Sawyer (2003, 2006), who developed the term ‘group flow’, states that it is possible to experience the flow as groups if the group has a common goal and every member of that group tries to attain it. However, it can be said that with routine activities and ordinary tasks in educational contexts, it may not be possible to raise students’ motivation level and guide them into currents, but recent studies on group flow suggest some solutions in this respect (Armstrong, 2008; Gaggioli, Milani, Mazzoni & Riva, 2011; Salanova, Rodríguez-Sánchez Schaufeli & Cifre, 2014).

Moreover, Dörnyei and Kubanyiova (2014) argue that it is possible to establish a group vision. In addition to these studies, there are other studies on group flow in educational contexts (Shernoff & Csikszentmihalyi, 2009; Shernoff,

Abdi, Anderson and Csikszentmihalyi, 2014). Shernoff (2013) argues that if the challenge and support level in an environment is high enough, it is possible to create group flow in that environment. Dörnyei et al. (2016) state that in classrooms, intensive group projects can be used to create DMCs opportunities for groups of students. According to them, group projects can be only be implemented when the goal is real and relevant to students' lives, there are tangible outcomes and when the criteria for successful behaviors are determined.

When the conditions required for individual DMCs are taken into consideration only, it is not possible to talk about group-DMCs. For group-DMCs to occur, the goals should be set clearly. In addition, in group projects whose aim is to make group-DMCs possible, there should be a sense of ownership and possession. That is, students should get the opportunities to take the responsibility of their own learning. In addition, the task challenges should be appropriate to students' skills levels. Another issue is that groups should be developed enough. That is, the group dynamics should be in balance and group members should accept and trust each other.

While talking about group-projects for groups DMC, it can be said that they can promote positive emotionality among groups. Deci and Ryan's (1985) self-determination theory supports this view.

The aforementioned theories and their sub theories so far can be used to explain several dimensions of Octalysis. Thanks to Self-Determination Theory, it is possible to provide a sound rationale for the CDs 2 (Development and Accomplishment), 3 (Empowerment of Creativity and Feedback) and 5 (Social Influence and Relatedness). By using DMC, it is possible to set a theoretical foundation for CDs 1 (Epic Winning and Meaning) and 4 (Ownership and Possession). In addition to these theories, research which focuses on the curiosity and motivation and approach and avoidance approach to goal attainment can be used.

Curiosity. From the early times, curiosity has received the attention of many researchers in terms of its relation with motivation. For example, researchers such as Arnold (1910) and Dewey (1913) examine interest-aided education by focusing on intrinsic lovely knowledge. In one of more recent studies, Oxford (2016) states

that curiosity can build up searching skills and thus should be taken into consideration while designing educational interventions. She also investigates the role of positive emotions which consist of happiness, curiosity, interest, pleasure and joy. She points out that these positive emotions are highly important in classrooms because they can broaden the individuals' attention, build toward innovative thought and actions, and contribute to resilience.

Silvia (2017) mentions the three strands of thought on curiosity's motivational nature. The first of these is that curiosity is a kind of deficit motivation. people get motivated when they realize a gap of knowledge and when they want to fill this gap. In addition, they may also want to get rid of the tension resulted from an unpleasant uncertainty and minimize the repellent effects of the drive. Secondly, curiosity is seen as a kind of intrinsic motivation. That is, it motivates people to explore and learn for their own sakes. Finally, there are studies focusing on the individual differences in curiosity. Because variation in curiosity can cause big differences in behavior and life outcomes.

When the historical development of curiosity is investigated, Silvia states that early studies (eg. Berlyne, 1960; Hull, 1943) purpose that curiosity and exploratory behavior are the results of the desires to reduce the novelty and uncertainty. By learning and exploring the world around, people can reduce something unpleasant such as uncertainty, ignorance, information deprivation and so on.

However, in 1970s the focus of research shifted from drive models to intrinsic motivation models. In this respect, intrinsic motivation which is free from extrinsic motivators such as rewards and punishments can be a source of inquiring and learning. According to Silvia (2017), this means that curiosity's motivating role is not related to people's desire to reduce something unpleasant.

When educational research is investigated, it is seen that there are studies on the use of curiosity (Hidi, 2001; Krapp, 2002; and Silvia, 2017). These studies argue that when classes get interesting, when students start to use complex reading and studying strategies for interesting materials and domains, and when they try to retain interesting material, they may get better grades. Similarly, Silvia (2017) states that curiosity can play a significant role to develop knowledge and competence.

To conclude, when Octalysis as a gamification model is investigated, it can be claimed that various theories can be used to clarify the theoretical foundations of the model. Self-Determination Theory, Flow Theory, Directed Motivational Currents Theory, Goal Theories, such as Approach and Avoidance Theory can be said to form the main pillars of the Octalysis model. In addition to these theories, it is also seen that the concepts of curiosity and scarcity can function within the model.

Chapter 3

Methodology

Research Method

In this study, a quasi-experimental mixed method was adopted (Creswell & Clark, 2011), and quantitative and qualitative data were collected in two phases. Figure 8 shows the general outline of the study in terms of the two phases and objectives, participants, data collection, and data analysis.

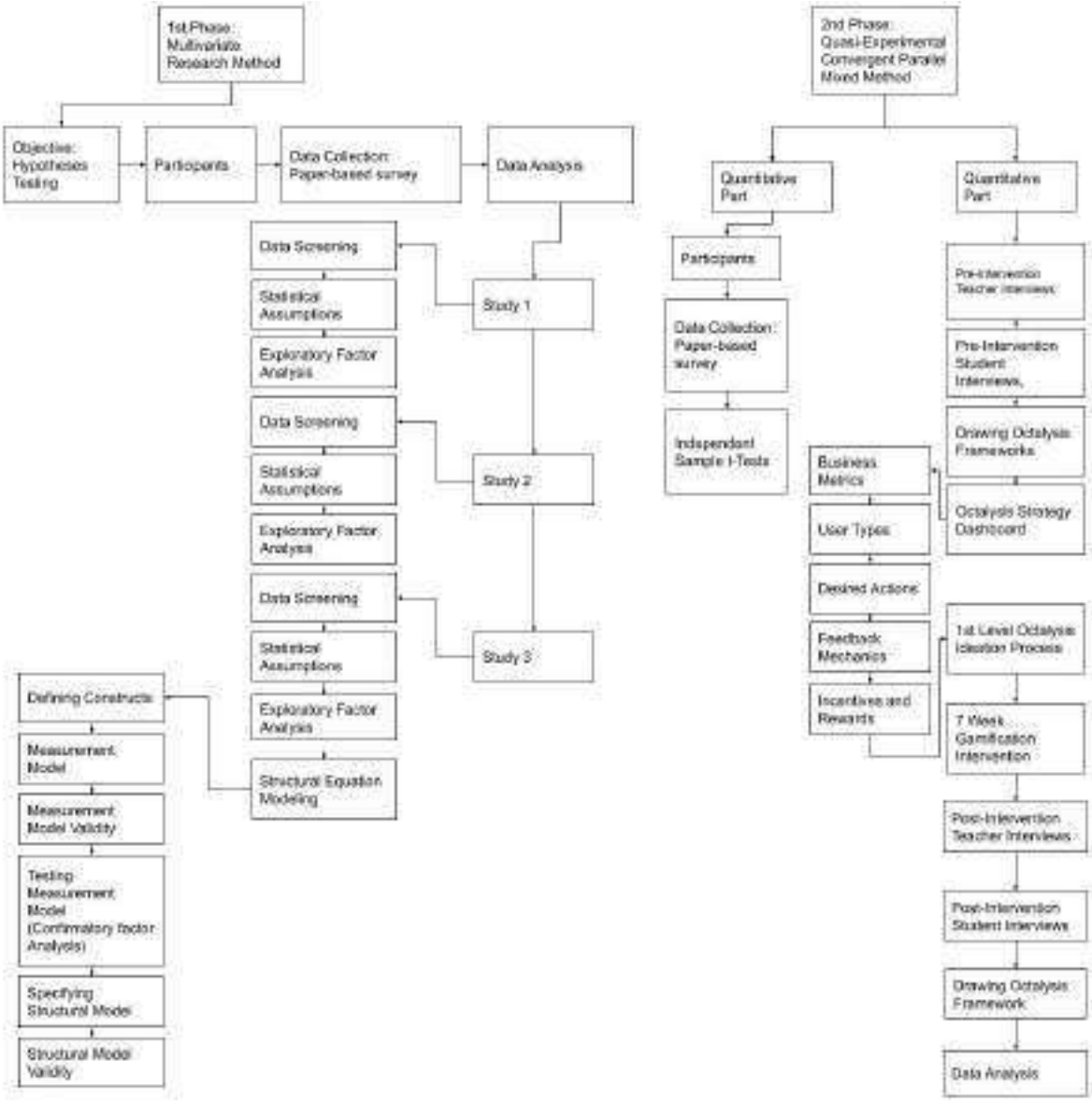


Figure 8. General outline of the study.

In the first phase of the study, 117 participants took part in Study 1. The number of participants in Study 2 was 337 whereas 1619 students participated in Study 3. The primary aim of the first phase of the study was to find out whether the surveys could be utilized in the way they had been proposed by their authors. In this respect, data were collected through paper-based administration of the scales which had been brought together in accordance with the theoretical foundation of Octalysis.

In each sub-study of the first phase, after data were collected, data screening procedures were applied to examine the suitability of the data sets for multivariate statistical tests. First, the type of missing data was determined. Then, the data sets were examined in terms of the extend of the missing data both per case and items. After that by conducting MCAR tests, the randomness of the missing data was checked. In line with the findings of MCAR tests, a suitable imputation method was administered. Finally, the data sets were inspected for the outliers.

When the data set in each sub-study was cleaned, the next step included examining the data set for any violation of statistical assumptions in terms of normality and linearity. Following the testing of statistical assumptions, exploratory factor analysis (EFA) procedures were applied. In scope of EFA procedures, the suitability of the data sets was controlled. Provided there were violations and the data were suitable for EFAs, separate EFA procedures were carried out for each different survey included in the study. Among these DMC, scarcity and psychological ownership dimensions consisted of one factor structures and tested in the same EFA. Then a different EFA was conducted to test Basic Psychological Needs Satisfaction and Frustration Scale consisted of three factors. Finally, separate EFAs were run for Achievement Goal Questionnaire and Curiosity and Exploration Inventory scales, each of which was consisted of two factors.

Following EFAs, a structural equation (SEM) modelling procedure was carried out to investigate whether it was possible to obtain a valid model. In the scope of SEM, first individual constructs were determined in accordance with the EFA findings. Then, a measurement model was created and its validity and reliability were checked. After that, by conducting a confirmatory factor analysis procedure, the measurement model was tested. According to the findings of CFA, a structural model was specified. Finally, the validity of the structural model was checked.

Having completed the analyses in the first phase of the study, in the second phase it was aimed to examine the practicality of the DMC integrated Octalysis gamification design in second language learning contexts. The second phase of the study consisted of two parts. In the first part which was a quantitative study a paper-based survey was carried out with the participation of 50 students in two classrooms which were named as experimental and control groups in Muğla Sıtkı Koçman University, School of Foreign Languages. Then, their responses were coded in SPSS and necessary data screening procedures were carried out. Having obtained cleaned data set, independent sample t-tests were carried out to investigate whether there was any significant difference between the two groups before starting the gamification intervention.

The second part of the second phase which lasted seven weeks provided both quantitative and qualitative data. First, pre-intervention interviews were carried out both with the teachers and students. In interviews, participants were asked to evaluate their language classes according to the Octalysis and then, they were asked to give a grade for each CD of the Octalysis model. The quantitative data obtained from the interviews and the pre-intervention survey were combined. Consequently, Octalysis frameworks were drawn for both groups and the situations in those groups were evaluated. After that, subsequent interview sessions were carried out to have the opinions of teachers and students about the writing classes. The grades given by teachers and students were combined and Octalysis frameworks were drawn.

The examination of the Octalysis frameworks provided information related to the weaknesses and strengths of both classes. In line with these Octalysis frameworks, an Octalysis dashboard was formed and a systematic gamification design was created. First, business metrics were identified and then user types were examined and desired actions were decided. After that, feedback mechanics, and incentives and rewards were determined to help students achieve desired actions. Finally, the gamification design was investigated by conducting an ideation process.

Having designed the gamification system, 7-week intervention was carried out. In the first week, students were informed about the system. During the following six weeks, the students took part in writing activities. In control group, the students followed routine procedures in their writing classes and no intervention was made

in terms of gamification. In the experimental group, all the writing activities were planned in accordance with the principles of Octalysis and findings obtained in pre-intervention interviews. Meanwhile, teachers were asked to keep regular records of their writing classes in terms of their students' motivation levels.

Following the intervention, post-intervention interviews were carried out with the participation of ten students from both groups. The findings obtained in these interviews were used to draw post-intervention Octalysis frameworks for both groups. Then, the findings were compared and analyzed.

Study Group

This study consisted of various sub-studies in each of which different groups of participants took part. In the first phase of the study which was comprised of three studies, a total of 2067 individuals from six universities (Izmir Institute of Technology, Erzurum Atatürk University, Muğla Sıtkı Koçman University, Aydın Adnan Menderes university, Burdur Mehmet Akif Ersoy University, and Çanakkale Onsekiz Mart University) across Turkey participated by completing the surveys in 2017-2018 and 2018-219 academic years. All these universities had School of Foreign Languages at the time of the study and all students were attending compulsory English language preparatory programs in their universities. Although there were some differences in terms of the way the language education given, all universities were following CEFR based language curriculums. Moreover, all students were having language courses at B2 level.

In the first study of the first phase, a total of 108 participants enrolled in eleven different programs in Muğla Sıtkı Koçman University (MSKU) responded the survey. 46 of these participants were female whereas 62 of them were male. The age range was between 18 and 29 with a mean age of 19.82, which meant they were almost at the same age. The data were collected in the second term of the 2017-2018 academic year. At the time of data collection, all the students had already been attending English as a foreign language compulsory preparation program in MSKU. In addition, they were all following a coursebook at B2 (upper-intermediate) level.

In the second study of the first phase, 338 participants from 34 different department in two different universities (Muğla Sıtkı Koçman University and Erzurum Atatürk University) responded to the survey in the second term of the 2017-

2018 academic year. 147 participants out of 338 were female and 170 of them were male. 21 participants did not prefer to state their sex. Their ages ranged between 18 and 33 with a mean age of 19.68. As in the first study, although there were a few participants over 25, the average score indicated that most of the participants were at the same age. The students who enrolled in MSKU were different from the first study. That is, the participants who had taken part in the first study were not included in the second study. In terms of their language proficiency levels, all the participants had been attending compulsory language preparation programs at B2 level in their universities for the past seven months.

Finally, the number of participants who handed their surveys and accepted to take part in the third study of the first phase was 1621, 755 of whom were female while 811 were male. 54 of the participants preferred not to state their sex. The participants were the students from 54 different departments of five different universities (Izmir Institute of Technology, Muğla Sıtkı Koçman University, Aydın Adnan Menderes university, Burdur Mehmet Akif Ersoy University and Çanakkale Onsekiz Mart University) in Turkey. The data were collected in the second term of the 2018-2019 academic year. As in the previous studies, the participants were attending compulsory language education programs in School of Foreign Languages in their universities and they were all following B2 level curriculums. Participants' ages varied between 18 and 32 with the mean age of 19.36.

In all the three studies, the students were asked to give their consents by signing the consent papers attached to their surveys. They were informed about the study and the procedure and told about their rights in the scope of the study. It was clearly stated that the data that would be collected through surveys would only be used for research purposes and they were free to leave the study anytime they wanted. In addition, email addresses and contact numbers of the researchers were provided in case they would like to learn about the results of the study or demand their data to be excluded from the study.

The second phase of the study consisted of two main parts. In the first part of the second phase of the study, 35 students participated (15 from the control group and 20 from the experimental group) in the survey in the 2018-2019 academic year. They were among the students who had not participated in any data collection processes in the first phase. All the students were attending the School of Foreign

Languages in Muğla Sıtkı Koçman University. They were having B2 level compulsory English language education and at the time of the study they were following an upper-intermediate (B2) level coursebook. That is, they were at the same level of English proficiency. Just like in the first phase, the students were informed about the procedure and they were told about their rights. Finally, they were asked to give their consents.

In the second part of the second phase, the students from experimental and control groups were invited to attend pre- and post-interviews to share their thoughts about the language education they were taking. Five students from the experimental and five students from the control groups responded positively.

These ten students were informed about the interviews. They were clearly told that interviews would be audio-recorded. They were free to demand to leave the study anytime they want and their audio-recordings would be deleted right away after they were transcribed. In addition, they were told that their recordings would not be shared by third parties and would be used for research purposes. Finally, they were asked to give their consents both in written and oral forms.

Finally, the teachers who had been carrying out classes in both groups were invited to take part in the interviews. Four teachers accepted the invitation (2 from experimental, 2 from control group). Like students, teachers were informed about the contents of the interviews, their rights in the scope of the study and voice-recording issue. Three of the teachers were teaching English as a foreign language more than 15 years and one teacher was at the beginning of his/her career with only two years of teaching experience. At the time of the study, two of the teachers were dealing with only routine pacing of the classes and they were mainly following the coursebook. The other two teachers (one in each classroom) were responsible for writing classes. Therefore, the post-intervention interviews were carried out with these two teachers.

Data Collection Process

All the data in the first phase collected by using the developed scale in which adapted versions of scales that were thought to reflect the eight core drives of Octalysis were included. The paper-based survey sheets had been posted to the universities that took part in the study and experts from the field of English as a

foreign language teaching helped the administration of the surveys. In order not to cause any misunderstanding, the scale was applied in Turkish since participants' English proficiency levels might not be enough to express their thoughts thoroughly. To ensure that items in the scale measured the same things with the original ones, translation and back translation method had been applied in advance (Erten, 2015). Having completed the translation process, the scale was revised and administered.

In the second phase, the quantitative data was collected by the researcher in paper-based format. As to the qualitative data, the researcher also carried out the interviews which were audio-recorded. Pre-intervention interviews were conducted between 21.02.2019 and 07.03.2019 and approximately five hours of recording was obtained. On the other hand, post-intervention interviews were carried out between 14.05.2019 and 17.05.2019, and yielded approximately three hours of recording. Moreover, in teachers from the experimental and control classrooms were asked to keep records of the unusual things they witnessed during the intervention.

Data Collection Tools

In accordance with the aim and method of the study, suitable data collection tools were selected. To collect the quantitative data, a scale was constructed by complying validated scales according to the theoretical background of the study. That is, each selected scale could be used to explain one or more CDs in Octalysis tool. Table 1 shows the relationship with the scales and the CDs. In the collection of the qualitative data, interviews and teacher logs were used.

Table 1

Selected Scales and Their Counterpart Core Drives in Octalysis

Scale	Corresponding CD(s)
The Curiosity and exploration inventory II	CD 7 – Unpredictability and Curiosity
	CD 2 – Development and Accomplishment
Basic psychological need satisfaction and frustration scale – general measure	CD 3 – Empowerment of Creativity and Feedback
	CD 5 – Social Influence and Relatedness
Achievement goal questionnaire revised	CD 8 – Loss and Avoidance
Directed motivational currents disposition questionnaire	CD 1 – Epic Meaning and Calling
Psychological ownership questionnaire	CD 4 – Ownership and Possession
Susceptibility to persuasion scale	CD 6 – Scarcity and impatience

The Curiosity and exploration inventory II (CEI-II). CEI II is a self-report instrument assessing individual differences in the recognition, pursuit, and integration of novel and challenging experiences and information. The Curiosity and Exploration Inventory II is developed by Kashdan, Gallagher, Silvia, Winterstein, Breen, Terhar and Steger (2009) to measure individual differences in broad dimensions of curiosity. It consists of ten five-point Likert scale items ranging from 1 (very slightly or not at all) to 5 (extremely) with two factors. Items 1, 3, 5, 7, and 9 reflect the first factor (stretching – motivation to seek out knowledge and new experiences) whereas items 2, 4, 6, 8, and 10 reflect the second factor (embracing – willingness to embrace the novel, uncertain and, unpredictable nature of everyday life).

CEI-II has good internal reliability (αs=.75-.85) (Kashdan et al., 2009; Ye, Ng, Yim, & Wang, 2015) and shows moderately large positive relationships with intrinsic motivation, reward sensitivity, openness to experience, and subjective vitality. Moreover, the CEI-II has shown incremental validity over and above the overlapping constructs of positive affect and reward sensitivity (Kashdan et al., 2009; Kashdan, McKnight, Fincham and Rose, 2011).

Basic psychological need satisfaction and frustration scale – general measure (BPNSF-General). BPNSF-General (Chen, Vansteenkiste, Beyers, Boone, Deci, Duriez, Lens, Matos, Mouratidis, Ryan, Sheldon, Soenens, Van Petegem, Van der Kaap-Deeder, & Verstuyf, 2015) consists of 24 five point Likert scale items which include a balanced combination of satisfaction and frustration items: Autonomy satisfaction (items 1,7,13 and19), autonomy frustration (items 2, 8, 14 and 20), competence satisfaction (items 5, 11, 17 and 23), competence frustration (items 6, 12, 18 and 24), and relatedness satisfaction (items 3, 9, 15 and 21) and relatedness frustration (items 4, 10, 16 and 22). Chen et al. (2015) state that each sub-scale has a Cronbach's alpha above .80. Hu and Zhang (2017) give similar reliability values (α > .80) for each subscale in their study which they carried out in an EFL context.

Achievement goal questionnaire revised (AGQ-R). This scale which was designed to assess achievement goals within a course-specific context is developed by Elliot and Murayama (2008). It has twelve items with five point Likert scale ranging from 1 (always false) to 5 (always true). It is possible to obtain data in terms

of mastery approach goal (items 1, 7 and 3), mastery avoidance goal (items 5, 11 and 9), performance approach goal (items 4, 2 and 8) and performance goal avoidance (items 12, 10 and 6).

As a result of her study in which she examines the generalizability of AGQ-R in both course specific and general academic contexts, Apostolou (2013) reports that the internal consistency of scores of all four measures of the AGQ–R have high reliabilities with Cronbach's alpha values ranging from .90 to .98.

Directed motivational currents disposition questionnaire (DMCDQ). This questionnaire is developed and validated by Muir (2016). It consists of a few items and questions among which 12 statements pose easy flow (8 items) and challenge (4 items) facets of DMC via a 5-point Likert type response format (completely disagree, disagree, to some extent agree, agree, and completely agree) (Ghanizadeh & Jahedizadeh, 2017). It aims at revealing answers to the questions of what proportion of people have experienced DMCs; who have experienced DMCs; and what their DMCs looked like. The questionnaire consists of ten items with five-Likert scale.

Ghanizadeh and Jahedizadeh (2017) state that the DMC Disposition Scale was demonstrated to have strong internal consistency (Cronbach's Alpha = .84)

Psychological Ownership Questionnaire (POQ). It has long been known that "feeling of ownership" plays a significant role in human attitudes and behavior (Van Dyne & Pierce, 2004). Astaryan, Slevitch, Larzelere, Morosan, and Kwun (2013) provide a scale by adapting the original scale by Van Dyne and Pierce for educational contexts. Accordingly, the scale consists of five statements regarding the psychological ownership to which participants responded on a five-point Likert scale from 1 (never) to 5 (very often). They state that the scale has strong internal consistency (Cronbach's alpha = .89).

Susceptibility to Persuasion Scale (STPS). Persuasive technologies are defined as a class of technologies that are intentionally designed to change a person's attitude or behavior. In this respect, Cialdini (2012) developed six principles of persuasion: reciprocation, scarcity, authority, commitment and consistency, consensus, and liking.

Following Cialdini, Kaptein, Ruyter, Markopoulos and Aarts (2012) have developed STPS which is an instrument that can measure users' susceptibility to persuasion. In the scope of current study, only the part related to the scarcity principle in the scale will be used. It has 5 statements with 7-point Likert scale. Borges, Isotani, Durelli, Reis, Bittencourt and Mizoguchi (2017) state that the scale has a high internal consistency (Cronbach's alpha = .82)

Interviews. Interviews made up the main data collection instrument of the second phase of the study. To get the participants' opinions about the gamification application in language classrooms and learn about their stories, the semi-structured interview technique was utilized with voluntary participants. All the sessions were voice recorded provided that the participants gave their consents and the data were analyzed by using the discourse analysis method. During the interviews both with teachers and students, not only qualitative data were collected but also quantitative data were gathered. At the end of each interview, the participant was asked to grade eight CDs of Octalysis. The quantitative data obtained from interviews were combined with survey results and they were used to calculate average scores for each CD which were to be used to draw Octalysis frameworks.

Teacher logbooks. To enable triangulation and provide more sound data, teachers were asked to write logbooks to record and keep track of their classrooms in terms of the application of gamification in classroom and students' reactions. According to van Meerkerk (2017), logbooks are a good way to capture the personal experiences of teachers especially when they are used to supplement the data from other sources.

Data Analysis

This study which was a quasi-experimental study composed of several quantitative and qualitative studies which were conducted at two main phases. Quantitative data obtained throughout the study were analyzed by using IBM Statistical Package for Social Sciences Version 21 (SPPS 21), which is a widely used software program for statistical analyses, and IBM SPSS Analysis of Moment Structures (AMOS) software which is a sub-module of SPSS software and which specifically used for structural equation modelling, path analysis and confirmatory

factor analysis. Both programs were provided by Hacettepe University, Turkey for students with community license. In addition, the qualitative data were analyzed by using discourse analysis method.

At the time of the study there were no known survey by the researcher to measure all dimensions of Octalysis. Therefore, before starting multivariate analyses, the Octalysis theory had been investigated thoroughly, and literature was scanned for pre-validated scales that could be used within the scope of this study. Six scales which were explained in detail in the data collection tools section were brought together and adapted in accordance with the research objectives. The items which made up Octalysis survey were translated into Turkish by an expert in second language teaching field. As Harkness and Schoua-Glusberg (1998) point out back-translation process was applied to check the suitability of the items. Then, the Turkish version of the survey was reviewed by another expert (a professor) in Turkish Language and Literature Department from Literature Faculty in MSKU. In accordance with the suggestions made by the experts, necessary adaptations were done. Finally, the survey was reviewed by two other experts who had been working in ELT field more than 15 years. The Survey first applied in Study 1 and then, with new adaptations in line with the findings, it was used in Study 2. Before final application of the survey in Study 3, it had been revised once more according to the findings of Study two.

Phase 1

In the first phase of the study, having completed the translation and application stages of the questionnaire, first preliminary statistical analysis and reliability analysis were conducted. Following the reliability analysis, EFA, CFA and SEM were conducted to obtain a reliable and valid scale. Figure 8 shows the process in this phase.

According to Holye (1995), this technique can be used to test the relationships between observed and latent variables. Teo, Tsai and Yang (2013) state that it is possible to use SEM for hypothesis testing, that is, for confirmatory aims. Ringdon (1998) points out that as a method, SEM can be used to represent, estimate and test theoretical network. SEM is carried out for mainly two reasons (Kline, 1998). It can be used to understand the patterns of correlations/covariance

among a set of variables. In addition, it can be used to explain as much of their variance as possible with the model specified.

For each EFA process in Phase 1 of the study, the same statistical procedures suggested by Hair et al. (2013) were applied. Prior to EFAs, first data screening processes proposed by Hair et al. (2013) were carried out to prepare the data sets to analyses. As the first step of the data screening processes, the data files were checked (Tabachnick & Fidell, 2013) in terms of missing values, and it was made sure that the missing values had not resulted from wrong data entry. Then, For the missing data diagnosis procedure, the four-step process suggested by Hair et al. (2013) was utilized at all the phases. Accordingly, this process includes the determination of the type of missing data, determination of the extent of missing data, diagnosing the randomness of the missing data process, and finally, selecting an imputation method.

To determine the extent of the missing data, the data sets were investigated for mean scores, standard deviations, and count and percent of the missing data per case and per variable were examined. In the next step of data screening procedure, the randomness of the missing data was controlled. For this purpose, Little's Missing Completely at Random (MCAR) test was used and depending on the findings an imputation method was selected. In study 1 and Study 2, estimated means suggested by MCAR test results were replaced with the missing values. However, since Study 3 was the main study, different imputation methods suggested in the literature were investigated. In this vein, an extra analyses was carried out in Study 3 and means mean scores and standard deviations of EM imputation, regression imputation and series modes were compared, and according to the findings it was decide to replace the missing values with series modes.

In the next step of the data screening procedure, the data sets were checked for the existence of outliers which have the possibility of affecting EFA analysis. At all the phases of the study, an outlier detection method was applied and how they were handled was explained. One of the reasons of having outliers could be the procedural errors (Tabachnick & Fidell, 2013). To find out such kind of outliers, exploratory analyses were run and frequencies of the data were examined. In case there could be extreme values out of the range, these outliers were fixed by controlling participants papers. Another type of outliers that could be seen might be

due to the unengaged responses by the participants. That is, if participants had given the same answer to all the questions, it could be an indication that they had completed the survey in an unengaged way. For this purpose, the standard deviations for all cases were calculted and essential changes were done.

Although as Hair et al. (2013) argue that the data set should be check for outliers from univariate, bivariate and multivariate perspectives, as Zijlstra (2009) points out in data sets which had been composed of Likert type statements checking these kinds of outliers might not be appropriate. Thus, no other outlier detection method was applied and proceeded for further analyses.

Having completed the procedures related to the missing data and outliers, the data sets were cleaned. After that, the next step included testing for the assumptions for multivariate data analysis since most multivariate analyses are based on several assumptions. Hair et al, (2013) point out that four of these assumptions have potential effects on every univariate and multivariate statistical techniques. In the scope of this study, due to the nature of the data sets, the two of them were checked: normality and linearity. (pp. 69-74). To that end, first the sample sizes which have a close relationship with normal distribution were controlled. Then, since the shape of the distribution of the data was another factor affecting the normality assumption, Skewness, standard error of Skewness, Kurtosis and standard error of Kurtosis values were examined (George & Mallery, 2013).

Having controlled the data in terms of normal distribution, data were also checked for linearity and any signal of curvilinearity was sought. For this purpose, scatter plots were used. However, since the number of the items used in the survey was high and potentially it was possible to create scatter plots over thousand, it was not found to be practical and scatter plots for only the items with the highest and lowest Skewness values (Tabcahnick and Fidell, 2013).

To sum up, with the above-mentioned analyses and investigations, data sets were prepared for EFA analyses which were explained below.

Then, all the data sets were subjected to ‘exploratory factor analysis’ (EFA) to see whether it was possible to use adapted versions of scales as suggested by their authors. Although all the scales and questionnaires were used in full versions in data collection phase, only the subscales that are related to Octalysis were factor

analyzed. Since sub-scales of different scales and questionnaires were compiled together and some of those consisted of second order structures, EFA procedure was applied separately for each subscale with second order structures (Mastery Avoidance and Performance Avoidance, Competence Satisfaction, Autonomy Satisfaction and Relatedness Satisfaction, and Stretching Curiosity and Embracing Curiosity). The scales which did not include any higher order factors (DMC, Scarcity and Ownership) were grouped together and then factor analyzed together. Table 2 shows the four different EFA procedures and the structures tested.

Table 2

Scales and Selected Dimensions to Explain Each CD of Octalysis in Each EFA

EFA	Name of Original Scale/Questionnaire	Related Sub-Dimension	Explained CD
1st EFA	Directed Motivational Currents Disposition Questionnaire	-	CD 1: Epic Meaning and Calling
	Psychological Ownership Questionnaire	-	CD 4: Ownership and Possession
	Susceptibility to Persuasion Scale	-	CD 6: Scarcity and Impatience
2nd EFA	Basic Psychological Needs Satisfaction and Frustration Scale -General Measure	Competence Satisfaction	CD 2: Development and Accomplishment
		Autonomy Satisfaction	CD 3: Empowerment of Creativity and Feedback
		Relatedness Satisfaction	CD 5: Social Influence and Relatedness
3rd EFA	The Curiosity and Exploration Inventory 2	Stretching Embracing	CD 7: Unpredictability and Curiosity
4th EFA	Achievement Goal Questionnaire-Revised	Mastery Avoidance Performance Avoidance	CD 8: Loss and Avoidance

For all EFA procedures, first the data sets were controlled regarding their suitability for factor analysis. To test whether they met the required assumptions, first sample sizes were examined and whether they were above the minimum absolute sample size. Then, correlations among variables were controlled (Pallant, 2011). To this end, by using SPSS 21 software, bi-variate correlations and their significance levels at .01 and .05 levels were obtained. By using Bartlett Test of Sphericity (Bartlett, 1954) the significances of all correlations were investigated. After that, the assumption of measure of sampling adequacy (MSA) was checked by investigating Kaiser-Meyer-Olkin Measure of Sampling Adequacy (Kaiser, 1974) value.

Provided the findings of the above-mentioned preliminary analyses yielded satisfactory results, factor analyses were with Maximum Likelihood (ML) extraction method. Although it was possible to attain factor structures according to the eigenvalues after ML extraction, since the number of factors were known, “a priori criterion” (Hair et al., 2013) was applied and factor numbers were pre-determined. Thus, factor analyses were stopped when the desired numbers of factors were achieved. ML extraction provided unrotated factor matrix with communalities for each item, and sum of squares and explained variance for each factor. In case there were significantly low communalities, the items with those values were eliminated. Then unrotated factor matrix was examined again. Provided it was not possible to get clear structures, in the next step rotation was applied.

Promax Kaiser Normalization rotation method was used for rotating the factors because it was thought that factors were correlated with each other (Tabachnick & Fidell, 2013). First, promax rotated pattern and structure matrices for full sets of variables were scanned. Based on the findings necessary eliminations were done and after each elimination factor analyses were respecified. After that, data in terms of promax rotated pattern and structure matrices, factor loadings, communalities, eigenvalues and percentage of the explained variance were studied. Within this context required actions were done.

After completing factor analyses in each step, the unidimensionality of the structures were controlled by examining the degree of item loadings in a specific structure. Provided there were unidimensional factor structures, the consistency of item-level errors (Gaskin, 2018) were scanned. That is the reliability of the structures were controlled. To this end, Cronbach’s Alpha values for each factor structure were calculated and studied. Following reliability check, the validity of the structures were examined. First, convergent validity checked by examining the correlations in a single factor because high correlations were the indicators of high convergent validity. For discriminant validity factors needed to be uncorrelated (Gaskin, 2018), so correlations among factors were scrutinized. Finally, after all four EFAs in each study were completed the survey was examined in terms of the existence of face validity (Hardesty & Bearden, 2004). For this purpose, views of participants and experts were taken and suggested adaptations were applied before using the survey in the next studies.

Having finalized the EFA procedures and obtaining clear factor structures, in the next stage Structural Equation Modelling (SEM) procedures which consisted of measurement model and structural model analyses were applied. To this end, six-stage procedure of SEM suggested by Hair et al. (2013) was used. Accordingly, first, individual constructs were described by defining individual items, their Turkish versions, their labels used in the analyses and explained constructs. Then, for each construct of Octalysis framework, working definitions of the constructs were done.

In the second step, the overall measurement model was developed and a visual diagram was drawn. In the model eight latent constructs were indicated by oval shapes. One-headed arrows were used to indicate the causal paths from the construct to the measured indicators. Then, by using two-headed arrows all constructs were correlated. And error terms for each indictor was added. In the next step, to produce empirical results a study was designed to test the measurement model. Following this step, the specification and estimation of the model was done and estimated parameters, error terms for indicators, and the total number of estimated parameters were identified.

After the model has been formed, the measurement model validity was assessed. To this end, the conformity between sample covariance matrix and population covariance matrix was checked. To confirm the model validity goodness-of-fitness (GoF) indices were studied. In this respect, first, chi-square (x^2) values, degrees of freedom and relative chi-square values were examined. Then, some absolute fit measures such as Goodness-of-fit (GFI) index, Root Mean Square Error of Approximation (RMSEA) with its 90% confidence interval values, Root Mean Square Residual (RMR) and Squared Root Mean Square Residual (SRMR) were examined. After that, Normed Fit Index (NFI), Comparative Fit Index (CFI) and Relative Fit Index (RFI) as incremental fit indices were investigated. Finally, two parsimony fit indices which were Adjusted Goodness-of-Fit Index (AGFI) and Parsimony Normed Fit Index (PNFI) were controlled (Collier, 2020). By analyzing the findings, decisions were made regarding the good fit of the model.

Having completed investigating the fitness of the model, the convergent validity of the factors was checked. High factor loadings were expected to be the indicators of convergent validity. So, unstandardized factor loadings, standard errors, t-values and p-values were calculated. In addition to convergent validity,

standardized factor loadings, average variance extracted and construct reliability values were calculated and judgement regarding the construct reliability were made. As another indicator of the convergent validity, Average Variance Extracted (AVE) values were calculated by adding all squared standardized factor loadings and diving them by the number of items (Hair et al., 2013). Since AVE values were too sensitive to reach any conclusions (Collier, 2020), Composite Reliability (CR) values which are the proportion of true variance to total variance were utilized for judgements.

Another factor affecting the construct validity is discriminant validity which is calculated with the shared variance between constructs (Fornell & Larcker, 1981). The correlations obtained between constructs were obtained by means of correlation matrix. Then, they were squared and the results were compared to the AVE values. Due to the drawbacks resulted from using strict AVE values, to determine the discriminant validity Collier (2020) offers using Heterotrait-Monotrait Ratio of Correlations (HTMT). Therefore before making any decisions in terms of discriminant validity, correlations within each construct (monotrait correlations) and correlations between constructs (heterotrait correlations) were obtained through correlation matrix. Then, they montrait and heterotrait correlations were added together and squared and a new correlation matrix was prepared with these data. The final matrix was used to make decisions in terms of the HTMT ratios and required judgements were made regarding the existence of discriminant validity.

Having examined the constructs in terms of convergent and discriminant values, other diagnostic cues were studied to see whether it was necessary to make any alterations in the model. Thus, residuals which are the difference between the observed and estimated covariance terms were calculated and the values were examined according to the published thresholds. Moreover, as the last diagnostic cue, Modification Indices (MI) for each parameter and for each factor loadings were scanned.

Eventually, after completing the above EFA and CFA procedures, it was time to proceed with the second part of SEM analysis which required the specification of the structural model which included a visual diagram of structural equations. In this respect, the findings from the CFA were transformed into the structural model by

replacing the double-headed arrows with one-headed arrows indicating hypothetical relationships and endogenous and exogenous constructs were determined.

In the fifth stage the structural model was specified. First, structural theory was defined and hypotheses to be tested placed. And finally in the sixth stage structural model validity was assessed. The Gof Indices such as chi-square, degrees of freedom, probability, relative chi-square value, GFI, RMSEA, 90% of RMSEA, RMR, SRMR as absolute fit indices; NFI, CFI and RFI as incremental fit indices; and AGFI and PNFI as parsimony measures were calculated and the obtained results were compared with the ones from CFA. In the next step, standardized estimated of the loadings and path coefficients were investigated. Consistency between the loading estimates of CFA and SEM analyzed and decisions regarding the existence of composite reliability were made. To determine the validity of the model the significance of path estimates were checked. Unstandardized estimates, standard errors, t-values, standardized estimates were calculated and whether hypotheses were supported or not decided. Finally, the SEM analysis and the first phase of the study were finalized by controlling diagnostic cues (residuals and MIs).

To help understand the processes and conducted analyses, Table 3 provides information about the statistical analyses and related research questions to be answered.

Table 3

Statistical Procedures in the 1st Phase and Targeted Research Questions

Procedure	Major steps	Unit of analysis	Research question
Data screening	Type of missing data Determination of the missing data	Frequency check	Is it possible to obtain appropriate data sets for statistical analyses
	Randomness of the missing data	MCAR test	
	Selecting imputation method	EM/Regression/series mode	
	Checking for outliers	Frequency check, Standard deviations	
	Checking for normality	Skewness and Kurtosis values	
	Checking for linearity	Scatter plots	
EFAs	Preliminary analyses	Sample size, bi-variate correlations, Bartlett's test, KMO value	Can preselected scales be used as suggested by their

		ML extraction, factor loadings, communalities, eigenvalues, percentage of explained variance	authors to measure the eight core drives of Octalysis?
	Factor analyses	Promax rotation, pattern and structure matrices, communalities, eigen values, percentage of explained variance; factor correlation matrix	
	Checking for reliability	Cronbach's alpha	
	Checking for convergent validity	Correlation matrix for each single factor	
	Checking for discriminant validity	Factor correlation matrix	
	Defining individual constructs	Defining scale type, item description, item labels, construct labels, conceptual construct definitions	
	Developing measurement model	Drawing visual diagram	
	Designing empirical study, drawing measurement model	Specification of estimated parameters	Is it possible to measure what is intended to measure with the adapted version of the survey?
	Assessing measurement model validity: model fit	Checking Gof statistics: Chi-square, df, CMIN/DF, GFI, RMSEA, RMR, SRMR, NFI, CFI, RFI, AGFI, PNFI	
	Checking convergent validity of the measurement model	Unstandardized estimates, standard error, t-values, p-values	
	Checking discriminant validity of the measurement model	Standardized factor loadings, AVE values, CR values, HTMT ratio of correlations	
CFA	Checking diagnostic cues	Standardized residuals, MIs of error terms, MIs of factor loadings	
			English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with: their endeavors to achieve big and meaningful goals. their feeling of accomplishment, skills development and overcoming challenges.
SEM	Specifying structural model	Definition of structural theory, hypotheses, drawing structural model	
	Assessing the structural model validity	Checking Gof statistics: Chi-square, df, CMIN/DF, GFI, RMSEA, RMR, SRMR, NFI, CFI, RFI, AGFI, PNFI	
		comparison of standardized factor loadings and construct reliabilities for structural and CFA models	

		their involvement in creative processes. their feeling of the ownership regarding their own learning processes. their feeling of being related to a social group in their learning environments. their desire to obtain scarce things and their perception of closing deadline. their being curious about the things in their learning environments which get their attention and their encounter with unpredictable things. their endeavors not to lose something and to avoid failure.
Hypothesis testing	Structural parameters estimates: structural relationships, unstandardized estimates, S.E, t-value, standardized estimate	

Phase 2

Second phase of the study incorporates both quantitative and qualitative data analyses processes. In the first stage of the second phase, the survey which had been factor analyzed was applied in two classrooms (experimental group and control group) in School of Foreign languages in MSKU with the participation of 35 students. The data was coded into SPSS 21 software program and data screening processes that had been previously applied in the first phase was applied. However, since there would not be any multivariate statistical analyses at this stage, the statistical assumptions to carry out such kind of analyses were not checked.

Having completed the data entry and data screening processes, mean scores were calculated by adding the items that were utilized for each single factor together. These mean scores were later used in to draw the Octalysis frameworks together with the data obtained from the pre-intervention audio-recorded interview. After that, as sampling groups were chosen randomly, independent t-test were done. To this end, first whether there were any violations of the assumptions was controlled. Independence of the observations were examined and the data set was scanned for outliers in terms of the existence of procedural outliers or extreme values. Then, the normal distribution of the data controlled (Pallant, 2011) by means of skewness

and Kurtosis values. Finally, t-tests were calculated. Mean scores for each variable, standard deviations, t-values, degrees of freedom, p-values and mean differences were investigated. Findings were used to make judgements in terms of the statistically significant differences between the two groups.

At this stage, also, pre-intervention audio-recorded and semi-structured interviews were carried out with the participation of ten students and four teachers from the two groups.

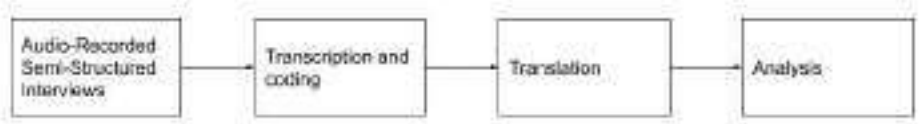


Figure 9. Research steps in the 2nd stage of the 2nd phase.

As Figure 9 shows the data collected through semi-structured and voice recorded interviews were first transcribed in accordance with Constant Comparison Grounded Theory (Glaser & Strauss, 1967). Grounded Theory aims at investigating any aspect of social world and tries to reach theories to explain how this works. The constant comparative method can be called as a method of data analysis to develop a grounded theory. Glaser and Strauss (1967) suggest that when used to generate theory, the comparative analytical method they describe can be applied to social units of any size. The constant comparative method involves breaking down the data into discrete 'incidents' (Glaser & Strauss, 1967) or 'units' (Lincoln & Guba, 1985) and coding them to categories.

Accordingly, data obtained from interviews were coded into categories which were labeled in accordance with the Octalysis framework. During the interviews, participants were asked to indicate their thoughts related to the general language education in their classrooms. Moreover, they were asked to name some of the most problematic issues. Finally, in each interview session, participants were asked to grade their general language education in terms of eight CDs of Octalysis on scale between 1 (minimum satisfaction) to 8 (maximum satisfaction).

The collected data were transcribed by the researcher and then, transcriptions were analyzed and they were categorized in accordance with the eight CDs of Octalysis. Throughout the analyses process similarities, differences and

variations of the data from the students and/or teachers from both groups were constantly compared (Lillemor & Hallberg, 2009). Having analyzed the initial interview data, participants were invited for the subsequent interviews to evaluate writing classes and state their opinions in terms of the efficiency of the activities in relation to the Octalysis framework. In addition, they were asked to grade as they did in the initial interviews. The same research steps were applied for the evaluation of writing classes.

Pre-intervention interviews not only provided insights about the issues in terms of both general language education and writing classes but also provided quantitative data which were used with the data from the pre-intervention survey. The mean scores attained from the survey, mean scores of the students and teachers were added together and new average scores were computed. These new scores were used to draw pre-intervention Octalysis frameworks for both general language education and writing classes. the findings were used the identify similarities and differences and determine the fundamentals of gamification intervention.

Based on the visual diagrams (Octalysis frameworks) that illustrated the roles of eight core drives in both classrooms, Octalysis strategy dashboard was prepared. To this end, first, desired actions were defined for four phases of engagement: discovery, onboarding, scaffolding and endgame. Then, feedback mechanics that were supposed to triggers those action were identified. In addition, business metric that were supposed to be used to measure the desired actions were put forward and incentives and win states were argued. Finally, level 1 Octalysis ideation process was carried out.

Having designed the intervention, for the later seven weeks the plan was implemented. In the first week students and teachers were informed about the process and the design in detail. Starting from the second week, gamification design was applied in writing classes in experimental group. Each week qualitative and quantitative data were collected by means of teacher's log book. The qualitative data was from the experience points given for tasks in each week. At the end of each week, experience points for each student in the experimental group was collected and students' engagement and interaction levels were pursued. The findings were later used to make comparisons within the experimental group and

between experimental and control group and graphically presented at the end of seventh week. In addition, throughout the intervention, copies of students' writings were collected and findings from these papers were compared with the quantitative and qualitative findings obtained from teacher's logbook.

Having completed the intervention process, students and teacher from both groups who had attended the pre-intervention interviews were again invited to take part in post-intervention interviews. As in the pre-intervention sessions, participants were asked to evaluate and grade the quality of their writing classes in accordance with the eight CDs of Octalysis framework. The recordings collected in post-intervention audio-recorded interviews were transcribed and categorized. Findings were used to identify any similarities and differences between the two groups. Finally, based on the gradings post-intervention Octalysis frameworks were drawn and these were compared with the pre-intervention Octalysis frameworks. The results were used to make decision regarding the applicability and practicality of DMC integrated Octalysis gamification intervention in English as a foreign language environments.

Table 4 provides a summary of the analyses and related research questions within the extend of second phase of the study.

Table 4

Stages in the Second Phase Analyses and Related Research Questions

Procedure	Major steps	Unit of analysis	Research question
Stage 1: quantitative part 1	Data screening	Visual inspection of the data set, series mode imputation method	Were there any differences between the experimental and control groups before the implementation of gamification intervention?
	Independent sample t-test	Checking preliminary assumptions, mean scores, standard deviations, t-values, df, p-values, mean difference	
Stage 1: qualitative part 1	Pre-intervention interviews with students from both groups, general evaluation	Transcription, coding, categorization, comparison of the findings, extracts as evidence for findings	Can semi-structured pre-intervention interviews provide data to draw Octalysis frameworks of experimental and control groups?
	Pre-intervention interviews with teachers from both groups, general evaluation		
	Pre-intervention interviews with students from both groups, writing evaluation		Is it possible to identify motivational drives by using pre-intervention Octalysis frameworks in experimental and control groups?
	Pre-intervention interviews with teachers from both groups, writing evaluation		
	Octalysis frameworks for general evaluation of students	Mean scores	

Stage 1: quantitative part 2	Octalysis frameworks for general evaluation of teachers	Mean scores	
	Octalysis frameworks for writing evaluation of students	Mean scores	
	Octalysis frameworks for writing evaluation of teachers	Mean scores	
Stage 2: quantitative part 1	Development charts for each week	Experience points	
Stage 2: qualitative part	Post-intervention interviews with students from both groups, writing evaluation	Transcription, coding, categorization, comparison of the findings, extracts as evidence for findings	How is the DMC integrated Octalysis gamification intervention perceived by the participants in the experimental group?
	Post-intervention interviews with teachers from both groups, writing evaluation	Transcription, coding, categorization, comparison of the findings, extracts as evidence for findings	
Stage 2: quantitative part 2	Octalysis frameworks for writing evaluation of students	Mean scores	Is there any difference between the experimental and control groups after the intervention?
	Octalysis frameworks for writing evaluation of teachers	Mean scores	

Chapter 4
Findings

Phase 1

This part of the study consisted of three different EFAs, one CFA and one SEM analyses.

Study 1

Data Screening

Type of missing data. In the first step, when the data in Study 1 were investigated, it could be seen that all the missing data were unknown. This means that missing data were not resulted from some errors that could occur at data entry stage but were caused by participants' not responding the items. As a result, the further examination of the data was carried out.

Determination of the extent of missing data. In the second step, which included the determination of the extent of the missing data, it was aimed at diagnosing the amount of missing data per case and per variable to plan the following steps (Hair et al., 2013). In the first data collection phase, which was Study 1, after omitting the demographic items which were not going to be used in the EFA, 117 variables were obtained for further examination of the data. When the missing values according to the variables were investigated, it could be seen that out of those 117 variables, 30 (25.64%) included at least one missing value. Table 5 shows the summary of missing data per variables for Study 1.

Table 5

Summary of Missing Data per Variable in Study 1

	N	Mean	Std. Deviation	Missing	
				Count	Percent
DMC5	105	2.87	1.421	3	2.8
Mastery Avoidance 1	105	3.84	1.001	3	2.8
Autonomy Frustration 3	105	2.82	1.262	3	2.8
DMC Length	106	1.77	1.382	2	1.9
DMC4	106	3.06	1.734	2	1.9
Performance Approach 2	106	3.82	1.076	2	1.9
Mastery Avoidance 2	106	3.94	.964	2	1.9
Competence Frustration 3	106	2.70	1.381	2	1.9
Belonging 1	106	2.97	1.268	2	1.9

Belonging 4	106	3.21	1.185	2	1.9
DMC Density	107	2.91	1.328	1	.9
DMC İntensity	107	2.68	1.371	1	.9
DMC7	107	2.61	1.503	1	.9
Mastery Approach 2	107	4.21	.898	1	.9
Autonomy Frustration 1	107	3.42	1.158	1	.9
Relatedness Frustration 1	107	1.90	1.072	1	.9
Competence Satisfaction 1	107	3.90	1.018	1	.9
Relatedness Satisfaction 2	107	3.71	.981	1	.9
Relatedness Satisfaction 3	107	3.69	.975	1	.9
Relatedness Frustration 3	107	1.97	1.145	1	.9
Competence Satisfaction 3	107	3.85	.979	1	.9
Competence Satisfaction 4	107	3.93	.924	1	.9
Embracing	107	2.36	1.144	1	.9
GLB Entity 1	107	2.86	1.489	1	.9
L2B Entity 3	107	2.44	1.422	1	.9
Psychological Ownership 1	107	2.93	1.257	1	.9
Psychological Ownership 2	107	3.05	1.152	1	.9
Psychological Ownership 4	107	2.68	1.051	1	.9
Psychological Ownership 5	107	2.79	1.180	1	.9
Belonging 3	107	2.81	1.167	1	.9

Although more than 25.64% of the variables had missing values, it could be seen from Table 5 that the total number of missing values was 43.

Table 6

Amount of Missing Data per Case in Study 1

Number of Missing Data per Case	Number of Cases	Percent
0	78	72.22
1	21	19.44
2	6	5.56
3	2	1.85
4	1	0.93
Total	108	100

Table 6 shows the count of missing values per case in Study 1. When it was examined, it was seen that most of the cases (72.22%) did not have any missing values. 19.44% of the 107 cases had only one missing value whereas 5.56% of them had two missing values. On the other hand, only two cases had three missing values. Finally, it could be seen that one of cases had four missing values. In sum, 27.78% of all cases had at least one missing value. Hair et al. (2013) state that if the number of missing data is under 10% for an individual case (p. 45), it can be disregarded on the condition that it is random. On the other hand, when the number of total values, which was 12,593, was taken into consideration, the number of missing values was only about 0.3% of the total, which meant there would not any

need to make changes at this stage. Thus, the next step of data screening procedure which was determining the randomness of the missing data was initiated.

Randomness of the missing data. Above findings suggested that the number of missing values both per variable and case was enough to proceed with further investigation of the data. The third step suggested by Hair et al. (2103) is determining the randomness of the data. For this purpose, Little's MCAR test was carried out. The finding showed that data did not have any patterns regarding the distribution of missing values. That is, data was missing completely at random (MCAR), Little's MCAR test: Chi-Square = 2673.858, DF = 2887, Sig = .998.

Selecting an imputation method. Accordingly, there were mainly two approaches to deal with the missing data as it was MCAR; imputation using only valid data and imputation by using replacement values (Hair et al., 2013, p.51). Since the former had some disadvantages that might cause a reduction in the sample size, the latter option, imputation by replacement values, was preferred. The finding of the Little's MCAR test suggested that estimated means (EM) were possible. So instead of deleting the cases with missing values, all the missing values were replaced with the estimated means from the above-mentioned test results.

Outliers. One of the reasons that cause outliers is the procedural errors (Tabachnick & Fidell, 2013). During the data entry process, at all the phases, each respondent's paper was given an identification code (ex. ID 123, 65, 198 etc.). In this way, it was possible to keep track of each paper. At all the phases, some items which seemed to have had errors and which were identified after running a frequency analysis were corrected. To find the correct entries, respondents' papers were referred by using the unique ID codes.

Another type of outlier may be due to unengaged responses by the participants. Therefore, the data were also examined in terms of unengaged responses by the participants. For this purpose, the standard deviation (SD) of all the values per participant was computed. If all the responses by an individual participant had been the same, the standard deviation of all that participants' responses would have been equal to zero. That is, for example, if any participant had answered all the items 'four' without reading the items, then the SD of them would have been equal to zero. When the standard deviation of the cases in total

was investigated, it was seen that none of the standard deviations for each case was too close to zero. In case of Study 1, the closest SD value was 0.8853, which meant that the participants replied items in an engaged way.

There are other reasons that might cause outliers to occur. Hair et al. (2013) suggest that data should be investigated in univariate, which means looking for extreme values for each variable; bivariate, which means extreme values between specific variable relationships; and multivariate, which is uncommon values for the model (p.69). But as Zijlstra (2009) stated, many of the proposed methods to detect the outliers are for continuous data and "cannot be used in multi-item questionnaire data" (p.89). When the nature of the data used in the scope of this study was considered, it was not appropriate to check for the outliers since the data set did not include any continuous variables except from age and achievement grades which were not used in the further analysis. In addition, all the variables were representatives of observations. Therefore, none of the outlier detection procedures were applied as Hair et al. (2013) suggested, and any possible outliers were retained to ensure the generalizability to the entire population (p. 65).

Statistical Assumptions

Normality. Perhaps the most influential assumption is normality. That is, the data need to be normally distributed across variables. When the individual variables are tested, it is called as univariate normality. On the other hand, when the combination of two or more variables are tested it is called multivariate normality. Hair et al. (2013) suggest that testing the univariate normality of all variables would be satisfactory in many cases (p. 69). The normal distribution of the data is based on both the sample size and the shape of the distribution. Hair et al. (2013) state that larger samples sizes decrease the negative consequences of nonnormality of the data whereas with small sample sizes, normality problems may have significant effects (p. 70). Regarding the sample size of the data collected in Study 1, although the sample size was not large enough, it had acceptable size.

Normal distribution of the data is also closely related to the shape of distribution, this can be tested by examining the measures of Skewness and Kurtosis. In other words, Skewness and Kurtosis values could be indicators of normality. Hair et al. (2013) state that if Skewness and Kurtosis values are zero, it

means that the data are normally distributed. Above or below zero indicates the violation of normality. There are researchers pointing out different values to be used as cutoff points. Çokluk, Şekercioğlu and Büyüköztürk (2012) indicate that Skewness and Kurtosis values between ±1 can be accepted as indicators of normal distribution. According to Hair et al. (2013), at .05 significance level, critical Skewness and Kurtosis values between ±1.96 indicate normal distribution. George and Mallery (2013) and Civelek (2017) obtain looser limit and state that Skewness and Kurtosis values between ±2 are also acceptable. Similarly, Sposito, Hand, and Skarpness (1983) argue that it should be between ±2.2. Even there are researchers who adopt more loose values such as Kline (2011) who points out that any Skewness and Kurtosis values between ±10 is acceptable. Based on these published thresholds, it could be said that the data achieved univariate normality. Table 7 shows the Skewness and Kurtosis values for the variables used in Study 1 in the scope of the current study.

Table 7

Skewness and Kurtosis Values of Study 1

	N		Skewness	Std. Error of Skewness	Kurtosis	Std. Error of Kurtosis
	Valid	Missing				
Mastery Avoidance 2	92	0	-1.478	.251	2.917	.498
Scarcity 5	92	0	-1.299	.251	.789	.498
Scarcity 3	92	0	-1.291	.251	.820	.498
Mastery Avoidance 1	92	0	-1.280	.251	1.671	.498
Performance Avoidance 3	92	0	-1.216	.251	1.568	.498
Relatedness Satisfaction 4	92	0	-1.189	.251	2.174	.498
Scarcity 2	92	0	-1.163	.251	.610	.498
Performance Avoidance 2	92	0	-1.078	.251	1.156	.498
Performance Avoidance 1	92	0	-1.073	.251	.457	.498
Competense Satisfaction 2	92	0	-.998	.251	1.127	.498
Mastery Avoidance 3	92	0	-.928	.251	.798	.498
Autonomy Satisfaction 1	92	0	-.909	.251	.372	.498
Competense Satisfaction 1	92	0	-.895	.251	.483	.498
Autonomy Satisfaction 3	92	0	-.862	.251	.856	.498
DMC12	92	0	-.854	.251	.400	.498
Relatedness Satisfaction 3	92	0	-.853	.251	.422	.498
DMC11	92	0	-.761	.251	-.073	.498
DMC1	92	0	-.734	.251	.251	.498
Autonomy Satisfaction 4	92	0	-.702	.251	.124	.498
Relatedness Satisfaction 2	92	0	-.648	.251	.442	.498
Competense Satisfaction 3	92	0	-.643	.251	.169	.498
Stretching 3	92	0	-.640	.251	-.172	.498
Competense Satisfaction 4	92	0	-.565	.251	-.373	.498
Relatedness Satisfaction 1	92	0	-.545	.251	.525	.498
DMC2	92	0	-.543	.251	-.382	.498
DMC4	92	0	-.537	.251	-.777	.498
Embracing 2	92	0	-.511	.251	.069	.498

Autonomy Satisfaction 2	92	0	-.506	.251	.254	.498
Embracing 5	92	0	-.504	.251	-.154	.498
DMC10	92	0	-.455	.251	-.642	.498
Stretching 1	92	0	-.355	.251	-.212	.498
Embracing 3	92	0	-.332	.251	-.425	.498
Stretching 5	92	0	-.312	.251	-.951	.498
Embracing 4	92	0	-.277	.251	-.568	.498
DMC6	92	0	-.267	.251	-.534	.498
DMC9	92	0	-.198	.251	-.530	.498
Psychological Ownership 2	92	0	-.192	.251	-.559	.498
DMC5	92	0	-.176	.251	-.300	.498
Psychological Ownership 5	92	0	-.157	.251	-.739	.498
Stretching 4	92	0	-.155	.251	-.300	.498
Psychological Ownership 1	92	0	-.148	.251	-.619	.498
DMC3	92	0	-.139	.251	-1.058	.498
Stretching 2	92	0	-.119	.251	-.419	.498
DMC8	92	0	-.060	.251	-.531	.498
DMC7	92	0	-.032	.251	-.744	.498
Scarcity 1	92	0	.013	.251	-1.040	.498
Psychological Ownership 3	92	0	.018	.251	-.821	.498
Psychological Ownership 4	92	0	.107	.251	-.081	.498
Embracing 1	92	0	.334	.251	-.713	.498
Scarcity 4	92	0	.343	.251	-.990	.498

When Table 7 is investigated, it can be seen that most of the Skewness and Kurtosis values of all the variables were between ±1 and others were between ±2 with only two exceptions which were Mastery Avoidance 2 (Kurtosis = 2.917) and Performance Satisfaction 4 (Kurtosis = 2.174). But they were not extreme values and it was decided to retain them. According to these findings, it was assumed that almost all variables were below the suggested thresholds and the data distributed normally and there was no violation of the normality assumption.

Linearity. The differences in Skewness values for variables indicate that there may be curvilinearity between some variables. Tabachnick and Fidell (2013) suggest using scatterplots to check the linear distribution of the data. However, with 50 variables, it was not practical to investigate all the scatter plots (over 1000) (Tabachnick & Fidell, 2013, p. 657). Thus, scatterplots of a few variable pairs were checked.

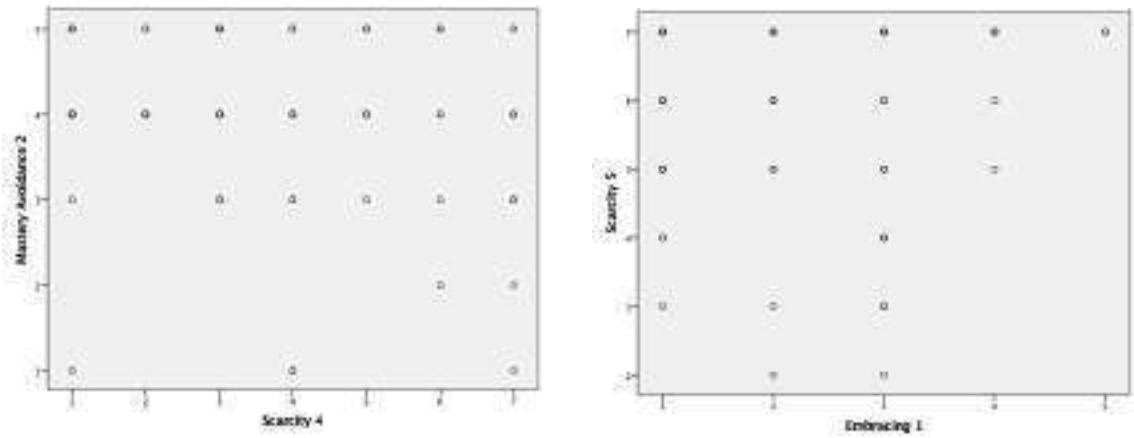


Figure 10. Scatterplots for linearity check among variables in Study 1.

Figure 10 shows sample scatterplots which have the highest negative (Mastery Avoidance 2 and Scarcity 5) and highest positive (Embracing 1 and Scarcity 4) Skewness values. Due to their Skewness values, these variables were expected to be among the worst. As can be seen from the plots, it is hard to talk about linear relationships among these variables but there is also not enough evidence that shows true curvilinearity. Therefore, no actions were taken regarding data transformation.

Exploratory Factor Analyses

The aim of the study was to test the Octalysis Theory (Chou, 2015). There was no known scale specially created for Octalysis at the time of application. Therefore, a scale had been created by adapting different parts of several other scales: The Curiosity and Exploration Inventory 2 (CEI-II) by Kashdan et al. (2009), Basic Psychological Need Satisfaction and Frustration Scale-General Measure (BPNSF-General) by Chen et al. (2015), Achievement Goal Questionnaire Revised (AGQ-R) by Elliot and Murayama (2008), Directed Motivational Currents Disposition Questionnaire (DMCDQ) by Muir (2016), Psychological Ownership Questionnaire (POQ) by Van Dyne and Pierce (2004), and Susceptibility to Persuasion Scale (STPS) by Cialdini (2012), all of which had been used, validated and published before. Although these scales and questionnaires were administered in their full format, only the sub scales which were related to the eight core drives of the Octalysis Theory were used in the analyses. That is, 50 out of 117 items were used in the factor analysis and these variables were supposed to be grouped into 10 factors.

Regarding the selection of type of factor analysis to be used, R-type factor analysis, which analyzes variable sets to find out the latent dimensions (Hair et al., 2013; p.96), was chosen.

Assumptions in Factor Analysis. After the omission of cases which had many missing values or consisted of unengaged responses, the sample size in Study 1 stage of the study was 108. In addition, cases which included zero DMC were also excluded from the analyses. That is, the responses of a participant were omitted if any kind of long-term motivation experience was not stated by that participant – in the scope of this study no experience of DMC. Thus, the data set with 92 sample size was obtained. Hair et al. (2013) state that the sample size to conduct a factor analysis should be larger than 100 and the minimum absolute sample size should not drop below 50. They add that in general it is demanded to have 5 to 1 ratio (at least five variables per variable) and more satisfactory ratio should be 10 to 1 (p. 100). When the number of variables which was 50 was taken into consideration, there had to be at least 250 observations to carry out factor analysis. In this sense, this could be seen as a problem. However, the sample size for Piloting stage was over the minimum number suggested by Hair et al. (2013). Therefore, the factor analysis was performed.

The 50 items had been compiled from the above-mentioned scales and questionnaires. These were supposed to group into 10 different factors. But, since these items were from different scales and questionnaires they were subjected to factor analysis separately. At first, the items for DMC, Scarcity and Psychological Ownership each of which was supposed to represent a single factor were subjected to factor analysis together. Then, items for BPNSF-General, CEI-II and AGQ-R were subjected to three different factor analyses. Each of these scales had at least two sub-sections and might have caused problems if they had been used as a single structure. In all factor analyses, Maximum Likelihood Extraction (MLE) Method which had been developed by Lawley in 1940s (Lawley & Maxwell, 1963) was preferred. MLE calculates loadings which maximize the probability of the observed correlation matrix and estimates population values for factor loadings (Tabachnick & Fidell, 2013, p. 641). This extraction method was preferred because it provides some advantages. In that, it is possible to compute indexes of the goodness of fit, it enables testing the statistical significance of factor loadings and calculating

correlations among factors (Cudeck & O'Dell, 1994). Also, it was the same extraction method which is used in AMOS software package which was going to be used in the Structural Equation Modelling (SEM) processes in the later stages of the study.

EFA with DMC, Scarcity and Psychological Ownership. The first factor analysis, assumptions for factor analysis were checked to see whether the data were suitable for factor analysis. Pallant (2011) suggests that to carry out factor analysis, the correlation matrix should include some correlations of r = .3 or greater (p. 187). When correlation matrix was examined, it was seen that although many correlation coefficients were below .3, there were some correlations of r = .3 or above some of which were significant at .01 level and .05 level. This indicated that it was possible to proceed the factor analysis. Another assumption to be met to proceed the factor analysis is that variables, some if not all, need to have significant correlations (Hair et al., 2013; p. 102). In this vein, to determine whether factor analysis is appropriate, Bartlett Test of Sphericity (Bartlett, 1954) was checked. Bartlett's test (approx. chi-square = 695. 613, df = 231), showed that correlations were significant at the .001 level when all the variables were taken into consideration. A third assumption to be met for factor analysis was the measure of sampling adequacy (MSA). The guideline provided by Hair et al. (2103) indicate that MSA value of ".80 or above, meritorious; .70 or above, middling; .60 or above, mediocre; .50 or above miserable; below .50, unacceptable" (p. 102). They add that to proceed factor analysis, .50 or above MSA value should be obtained whereas Kaiser (1974) state it should be above .60. The Kaiser-Meyer-Olkin measure of sampling adequacy value obtained in the analysis was .684, which was not satisfactory enough but above the recommended value. Thus, it was assumed that although the data were not perfect for factor analysis, it met the required assumptions to proceed.

Factor analysis with maximum likelihood extraction method yielded 7 possible factors with eigenvalues above 1. However, since the number of factors had already been known prior to factor analysis, "a priori criterion" (Hair et al., 2013; p. 107) was applied to determine the number of factors to be extracted. Hair et al., (2013) state that this approach can be used while testing a theory or hypothesis. Thus, factor extraction was stopped when three factors were obtained.

Table 8

Unrotated Factor Matrix for DMC, Psychological Ownership and Scarcity in Study 1

	Factor			Communality
	1	2	3	
DMC1	.411	-.009	.578	.503
DMC2	.377	.030	.486	.379
DMC3	.094	.269	.370	.218
DMC4	-.019	-.226	.022	.052
DMC5	.268	.075	.534	.362
DMC6	.408	.241	.475	.451
DMC7	.999	-.001	-.002	.999
DMC8	.479	.023	.176	.261
DMC9	.494	-.124	.158	.285
DMC10	.419	-.105	.254	.251
DMC11	.375	-.062	.479	.374
DMC12	.295	.093	.383	.242
Scarcity 1	.055	.252	.062	.070
Scarcity 2	.021	.209	.333	.155
Scarcity 3	.005	.212	.301	.136
Scarcity 4	.370	.177	-.218	.216
Scarcity 5	.199	-.025	.156	.065
Psychological Ownership 1	.062	.917	-.028	.846
Psychological Ownership 2	.151	.811	-.128	.697
Psychological Ownership 3	.112	.513	.023	.276
Psychological Ownership 4	-.094	.335	-.196	.159
Psychological Ownership 5	.021	.767	.052	.592
				Total
Sum of Squares (eigenvalues)	4.440	3.177	1.774	9.391
Explained Variance	20.180	14.443	8.063	42.686

Extraction Method: Maximum Likelihood.

As can be seen from Table 8, eigenvalues (sum of squares) for three factors were 4.440, 3.117 and 1.774 respectively and 9.391 in total. And these three factors explained 42.686% of the total variance without the application of any rotation methods. When the factor loadings in unrotated factor matrix were investigated, it was not possible to interpret these unrotated loadings. Therefore, as the next step, it was decided to proceed the rotation process. But, before that, the communalities, sum of squared factor loadings, were examined. Communalities are the values that indicate the level of correlation of an item with others. Hair et al. (2013) suggest that communalities should not be lower than .5 (p. 134). Osborne, Costello, and Kellow (2008) suggest that the cut-off point can be as low as .4 whereas Child (2006) argues that communalities below .2 should be deleted. Thus, items (DMC 4, Scarcity 1 and 5) that had the lowest communalities were deleted and factor analysis was respecified. Communalities were checked again and another item (Psychological Ownership 4) that had a low communality value was excluded, too. After that, it was

seen that the total variance explained by the three unrotated factors increased to 48.824%. Although items of scarcity had low levels of communalities, they weren't excluded not to lose the entire scarcity sub-section. Having eliminated the items with low communalities, it could be seen that there were still some items with communality values below the suggested cut-off points even though the elimination of four items resulted in increases in other communalities. However, since the percentage of total variance was increased considerably, factor analysis was not terminated.

In the next step, The Promax Kaiser Normalization rotation was carried out to obtain a clean set of factor loadings. As it is known, there are basically two main categories of rotation methods: orthogonal and oblique. According to Tabachnick and Fidell (2013), orthogonal rotation methods are used when it is thought that the factors are uncorrelated (p.642). On the other hand, oblique rotation is used when there are correlations among factors. In the scope of this study, since it was hypothesized that factors were correlated with each other, Promax rotation was preferred. After the rotation of the factor loadings was investigated, it was seen that all the loadings were found to be above .3. Tabachnick and Fidell (2013) state that as a general rule, only the factor loadings above .32 should be interpreted (p. 654). On the other hand, Hair et al. (2013) point out that in the identification of significant factor loadings, sample size plays highly important role. In a sample size of 50, a factor loading of .75 is required while with bigger sample sizes (350 or more), .30 factor loading can be accepted as significant. In this respect, it can be said that all the factor loadings could be accepted as significant as they are above .3. However, although Promax rotation improved factor loadings and provided clearer structures, factors were not free from some problems. The elimination of DMC 4, Scarcity 1 and Scarcity 5 due to their poor communality values caused DMC 7 to have an extremely high factor loading (1.017). In the third respecification of the factor model, DMC 3 and DMC 11 were deleted because of insignificant loadings. In the fourth run, DMC 8 and Scarcity 4 were deleted. In the subsequent turns, DMC 2, because of the low factor loading, and DMC 10 and Psychological Ownership 3, because of low communalities, were deleted. Finally, after all, a simplified factor structure was obtained. Tables 9 and 10 provide pattern and structure matrices which were

obtained after Promax rotation for full and reduced sets of variables together with the eigenvalues of factors and the percentage of total variance each factor explains.

Table 9

Promax-Rotated Pattern and Structure Matrices for DMC, Psychological Ownership and Scarcity: Full set of variables in Study 1

	Pattern Loadings*			Structure Loadings			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
DMC1	-.093	.634	.134	-.050	.692	.443	.503
DMC2	-.039	.535	.142	-.003	.601	.402	.379
DMC3	.199	.449	-.115	.233	.408	.100	.218
DMC4	-.228	-.001	-.008	-.228	-.021	-.005	.052
DMC5	-.012	.602	.002	.031	.602	.294	.362
DMC6	.172	.547	.158	.208	.636	.421	.451
DMC7	.068	-.062	1.027	.046	.441	.996	.999
DMC8	.021	.175	.402	.027	.371	.486	.261
DMC9	-.119	.136	.441	-.117	.341	.509	.285
DMC10	-.124	.252	.314	-.111	.396	.439	.251
DMC11	-.129	.516	.153	-.095	.581	.405	.374
DMC12	.038	.430	.103	.066	.483	.311	.242
Scarcity 1	.240	.098		.247	.115	.043	.070
Scarcity 2	.143	.404	-.165	.174	.334	.028	.155
Scarcity 3	.150	.369	-.166	.179	.299	.010	.136
Scarcity 4	.242	-.249	.471	.216	-.004	.346	.216
Scarcity 5	-.042	.162	.129	-.032	.222	.209	.065
Psychological Ownership 1	.911	.078	-.015	.917	.135	.007	.846
Psychological Ownership 2	.832	-.055	.136	.826	.071	.096	.697
Psychological Ownership 3	.507	.082	.052	.512	.143	.083	.276
Psychological Ownership 4	.361	-.177	-.033	.349	-.167	-.125	.159
Psychological Ownership 5	.745	.153	-.082	.758	.166	-.021	.592
							Total
Sum of Squares (eigenvalue)	2.663	2.874	2.051				7.588
Percentage of Explained Variance	12.105	13.064	9.322				34.491

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

Table 10

Promax-Rotated Pattern and Structure Matrices for DMC, Psychological Ownership and Scarcity: Reduced set of variables in Study 1

	Pattern Matrix*			Structure Matrix			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
DMC1	-.091	.647	-.025	.025	.622	.149	.396
DMC5	-.009	.592	-.051	.091	.575	.125	.333
DMC6	.112	.737	.035	.257	.769	.279	.605
DMC12	-.023	.585	.076	.102	.603	.246	.369
Scarcity 2	-.043	-.018	1.013	.170	.278	.998	.999
Scarcity 3	.095	.041	.407	.190	.181	.440	.205
Psychological Ownership 1	.918	-.050	.032	.915	.131	.213	.840
Psychological Ownership 2	.844	-.080	.024	.834	.085	.180	.701
Psychological Ownership 5	.745	.117	-.030	.760	.247	.164	.591

					Total
Sum of Squares (eigenvalue)	1.413	2.127	1.499		5.039
Percentage of Explained Variance	15.698	23.631	16.661		55.990

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

Table 11

Promax-Rotated Factor Correlation Matrix for DMC, Psychological Ownership and Scarcity in Study 1

Factor	1	2	3
1	1.000	.187	.214
2	.187	1.000	.300
3	.214	.300	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

When Tables 9 and 10 are examined, it can be seen that the factor loadings and communality values were improved after the elimination of problematic variables. But as can be seen in Table 10, Factor 3 (Scarcity) still had problems. First of all, only two variables remained. And it means that these scales cannot be used in SEM analysis because the scale will be underidentified. In addition, Scarcity 2 still had an extreme value which indicates another problem. On the other hand, Factor 1 (Psychological Ownership) and Factor 2 (DMC), it could be stated that optimal structures existed with all variables with high loadings on a single factor. Based on this finding, it could be stated that the final structures of Factor 1 and Factor 2 were unidimensional. That is, there were strong associations between each item and it was possible to obtain a single concept (Hair et al., 2013; p. 123). To be able to talk about the existence of unidimensionality, there should be highly loaded items on a single factor.

Since Factors 1 and 2 were unidimensional, the next step was to examine their reliability. According to Gaskin (2018), for a single factor, the consistency of item-level errors determines the reliability of that factor. Hair et al. (2013) point out that one way of assessing the reliability of factors is to check their Cronbach's alpha scores. They say that a threshold of .70 or above should be adopted although in some exploratory research it is possible to use .60 level (p. 125). In Study 1, the Cronbach's alpha for Factors 1 (Psychological Ownership), 2 (DMC) and 3 (Scarcity) were, .872, .731 and .606, respectively. As expected, the dimension of

Scarcity seemed to have issues regarding its reliability whereas Psychological Ownership and DMC dimension had reliability values at essential levels.

The next step after checking the reliability of factors was to assess their validities. That is, to what extent the factors represent what is intended. Hair et al. (2013) state that a scale needs to have face, convergent, discriminant and nomological validities (p. 124). Similarly, Gaskin (2018) points out that factors should have face, convergent and discriminant validities. Hardesty and Bearden (2004) say that face validity is related to the judgments of users, experts, participants about the use of a scale in terms of its appropriateness to the targeted aims (p.99). At this stage, face validity is not checked and left for the conclusion of EFA procedures in Study 1. Regarding the convergent validity, Hair et al. (2013) point out that it "assesses the degree to which two measures of the same concept are correlated" (p. 124). High correlations in a single factor mean that factor has convergent validity. When Table 10 is checked, it is seen that item loadings for Factors 1, 2 and 3 are sufficiently high to assume there was convergent validity for all the three factors. Finally, to assume the existence of discriminant validity, factors need to be uncorrelated and distinct (Gaskin, 2018). Farrell and Rudd (2009) point out that to talk about discriminant validity, a latent variable should be discriminated from other latent variables. Hair et al. (2013) point out that for testing discriminant validity, correlations among factors need to be checked. They add that low correlations show that factors are different from each other. When the Factor Correlation matrix provided in Table 11 is investigated, it can be seen that factors have low correlations (.187, .214, .300), which indicates that discriminant validity existed and factors were different from each other.

EFA with Basic Psychological Needs Satisfaction and Frustration Scale. For the second EFA analysis of Study 1, the same procedures were applied as in the previous part. This time, subdimensions of Basic Psychological Needs Satisfaction and Frustration in General Scale (Competence Satisfaction, Relatedness Satisfaction and Autonomy Satisfaction) were tested. First the data set was examined to check whether it was appropriate for EFA. For the first assumption, the correlation matrix was examined. Findings showed that there were many correlations of r = .3 or above, and out of 78 correlations, 33 were significant at .01 level (2-tailed) and 10 were significant at .05 level (2-tailed). This enabled going on

with the factor analysis. Then, Bartlett Test of Sphericity was checked and it was found that Bartlett test was significant at .01 level, chi-square = 308.953, df = 66, $p < .001$, which also meant that correlations were significant. As another assumption, measure of sampling adequacy was controlled and Kaiser-Meyer-Olkin measure of sampling adequacy value was found as .753, which was adequate according to the guidelines suggested by Hair et al. (2013; p. 102). Therefore, it was assumed that proceeding with factor analysis was suitable.

Factor analysis with maximum likelihood extraction method was applied. When the eigenvalues were examined, it was seen that there were 3 factors with eigenvalues above 1. Table 12 gives unrotated factor loadings, eigenvalues and percentages of explained total variance by each factor. It could be seen that eigenvalues for each factor was: Factor 1 = 4.127, Factor 2 = 1.485, and Factor 3 = 1.020. These three factors were explaining 55.270% of the total variance before rotation was carried out. When the communalities were investigated, it was realized that some values were below the suggested limit .4 (Osborne, Costello, & Kellow, 2008) or .2 (Child, 2006). Thus, items with the lowest communalities, Autonomy Satisfaction 1 (.182) and Relatedness Satisfaction 2 (.140) were eliminated before continuing further analysis.

Table 12

Unrotated Factor Matrix for BPNSF in Study 1

	Factor			Communality
	1	2	3	
Autonomy Satisfaction 1	.239	.353	-.007	.182
Autonomy Satisfaction 2	.306	.446	.012	.292
Autonomy Satisfaction 3	.383	.663	.014	.586
Autonomy Satisfaction 4	.241	.643	-.466	.689
Relatedness Satisfaction 1	.261	.310	.340	.280
Relatedness Satisfaction 2	.172	.246	.223	.140
Relatedness Satisfaction 3	.389	.513	.454	.621
Relatedness Satisfaction 4	.102	.365	.477	.371
Competence Satisfaction 1	.423	.264	-.106	.260
Competence Satisfaction 2	.498	.375	-.028	.390
Competence Satisfaction 3	.537	.294	-.218	.422
Competence Satisfaction 4	.999	-.003	.000	.999
				Total
Sum of Squares (eigenvalues)	4.127	1.485	1.020	6.632
Explained Variance	34.392	12.379	8.500	55.270

Extraction Method: Maximum Likelihood.

After the elimination of items with low communality values, unrotated total variance explained by three factors was increased to 61.422%. However, when

those items were deleted, the eigenvalue for Factor 3 dropped below 1. Thus, it was decided to retain those two items. When the factor loadings were investigated, it was hard to decide the factor structures. Although some items strongly loaded to a single factor, there were some cross-loaded items. For example, Relatedness Satisfaction 1 loaded both on Factor 2 and Factor 3, Competence Satisfaction 2 loaded on both Factor 1 and Factor 2, and Relatedness Satisfaction loaded on the three factors. Therefore, Promax Kaiser Normalization method was applied.

After promax rotation, it was seen that loading of Competence Satisfaction 1 was below .3. Therefore, the item was eliminated and factor analysis was rerun. Then, to reach a simple factor structure, Autonomy Satisfaction 3 and Competence Satisfaction 3 which had cross loadings were deleted and factor analysis was run again. Finally, Promax rotation revealed a simple factor structure. Table 13 provides the pattern and structure matrices for the full set of variables, eigenvalues and percentage of explained variance after rotation whereas as Table 14 gives pattern and structure matrices for the reduced set of variables, their eigenvalues and percentage of explained variance by three factors after Promax rotation. Accordingly, communalities and factors loadings improved considerably after eliminating problematic variables. It could be said that optimal structures were obtained because most of the variables strongly loaded on a single factor. It seemed that one of the major problems was that the dimension of Competence Satisfaction had two variables, which meant that like Scarcity this factor would be underidentified in SEM.

Table 13

Promax-Rotated Pattern and Structure Matrices for BPNSF in Study 1: Full set of variables

	Pattern Matrix*			Structure Matrix*			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
Autonomy Satisfaction 1	.301	.154	.039	.400	.325	.274	.182
Autonomy Satisfaction 2	.363	.218	.051	.500	.425	.350	.292
Autonomy Satisfaction 3	.551	.321	.000	.714	.600	.449	.586
Autonomy Satisfaction 4	.962	-.215	-.101	.799	.223	.315	.689
Relatedness Satisfaction 1	-.038	.518	.060	.255	.527	.284	.280
Relatedness Satisfaction 2	.014	.359	.015	.203	.373	.192	.140
Relatedness Satisfaction 3	.035	.737	.063	.441	.785	.431	.621
Relatedness Satisfaction 4	-.087	.699	-.155	.182	.582	.129	.371

Competence Satisfaction 1	.286	-.003	.298	.444	.282	.450	.260
Competence Satisfaction 2	.312	.135	.303	.543	.436	.534	.390
Competence Satisfaction 3	.396	-.116	.410	.558	.279	.569	.422
Competence Satisfaction 4	-.114	-.029	1.069	.447	.420	.994	.999
							Total
Sum of Squares (eigenvalues)	2.330	2.025	.876				5.231
Explained Variance	19.413	16.876	7.303				43.592

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

Table 14

Promax-Rotated Pattern and Structure Matrices for BPNSF in Study 1: Reduced set of variables

	Pattern Matrix*			Structure Matrix*			
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Communality
Autonomy Satisfaction 2	.156	.151	.360	.413	.479	.530	.328
Autonomy Satisfaction 3	.296	.086	.524	.589	.604	.716	.601
Autonomy Satisfaction 4	-.148	-.063	.883	.221	.422	.775	.625
Relatedness Satisfaction 1	.506	.007	.036	.526	.334	.273	.278
Relatedness Satisfaction 2	.395	-.074	.076	.385	.212	.209	.152
Relatedness Satisfaction 3	.621	.282	-.061	.762	.616	.408	.622
Relatedness Satisfaction 4	.779	-.170	-.109	.626	.227	.138	.434
Competence Satisfaction 2	-.141	.803	.046	.362	.748	.502	.572
Competence Satisfaction 4	-.030	.724	-.063	.377	.665	.393	.446
							Total
Sum of Squares (eigenvalues)	2.869	.771	.418				4.058
Explained Variance	31.877	8.570	4.644				45.091

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

When Table 14 is examined, it can be seen that unidimesionality of the factors was achieved since all variables loaded on a single factor.

The cronbach-alpha values for the reliability of each factor were: .658 for Factor 1 (Relatedness Satisfaction), .662 for Factor 2 (Competence Satisfaction), and .699 for Factor 3 (Autonomy Satisfaction), respectively. According to Hair et al. (2013) these figures were below the desired level of reliability, which is .7. However, since these values were close to .7, it could be said that these dimensions could be accepted as reliable. Since all the variables were sufficiently high, it could be assumed that convergent validity was achieved. Regarding the discriminant validity, Table 15 shows the intercorrelations of the factors.

Table 15

Promax-Rotated Factor Correlation Matrix for BPNSF-General in Study 1

Factor	1	2	3
1	1.000	.601	.460
2	.601	1.000	.649
3	.460	.649	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

Although correlations among factors were relatively high (.601, .460, .649), they were all below the threshold .7 which was suggested by Gaskin (2018). The results of this analysis support the use of subdimensions (Autonomy Satisfaction, Competence Satisfaction and Relatedness Satisfaction) of BPNSF as separate scales as suggested by the scale authors (Chen et al., 2015).

EFA with The Curiosity and Exploration Inventory 2. The third EFA analysis in Study 1 included the items assessing the curiosity within two factors: stretching and embracing. Before factor analysis, the suitability of the data set for EFA was checked by controlling whether the assumptions were met. When the correlation table for the items was examined, it was seen that there were many correlations of .3 or above. 23 of the total 53 correlations were significant at .01 (2-tailed) level whereas 6 of them were significant at .05 (2-tailed) level. This provides the rationale to proceed the factor analysis with these variables. Bartlett's Sphericity test (chi-square = 239.390, df = 45, $p < .01$) suggests that these correlations were significant. In terms of measure of sampling adequacy, Kaiser-Meyer-Olkin measure of sampling adequacy value was found to be .799, which was appropriate to assume that EFA could be carried out.

Maximum likelihood extraction method yielded three factors with eigenvalues greater than 1 but since it was known that there should be 2 factors, a priori criterion (Hair et al., 2013) was applied and factor extraction was terminated when 2 factors were reached. Table 16 shows the unrotated factor loadings, communalities, eigenvalues and percentage of explained variance by the two factors before the application of rotation. Accordingly, Factor 1 had 3.581 eigenvalue and explained 36.508% of the total variance while Factor 2 had eigenvalue of 1.303 and explained 13.031% of the variance. The two factors together accounted for 49.539% of the variance and their sum of squared loading was 4.954. Then the factor loadings were examined. Table 16 revealed that it was not possible to identify the factors without

rotation. It also showed that there were communalities with low values. Thus, before proceeding for further analysis the items (Embracing 1 and Embracing 5) were excluded.

When Promax rotation was applied, some improvements in factor loadings were observed.

Table 16

Unrotated Factor Matrix for CEI-II in Study 1

	Factor		Communality
	1	2	
Stretching 1	.365	.531	.415
Stretching 2	.538	.030	.291
Stretching 3	.645	.253	.479
Stretching 4	.642	-.195	.450
Stretching 5	.697	.131	.503
Embracing 1	.055	.027	.004
Embracing 2	.629	.121	.410
Embracing 3	.659	-.296	.521
Embracing 4	.694	-.268	.553
Embracing 5	.230	.259	.120
			Total
Sum of Squares (eigenvalues)	3.651	1.303	4.954
Explained Variance	36.508	13.031	49.539

Extraction Method: Maximum Likelihood.

But there were still some problems because it was still not possible to reach a simple two-factor structure. Stretching 2, Stretching 4, Stretching 5, and Embracing 3 and Embracing 4 strongly loaded on Factor 1 while Stretching 1, Stretching 3 and Embracing 2 loaded on Factor 2. Therefore, some items (Stretching 2 and Stretching 4) were excluded and the analysis was run again. Although the improvements were observed in the model, it was not possible to reveal clear structures. Table 17 shows the pattern and structure matrices after the respecification of the model.

Table 17

Promax-Rotated Pattern and Structure Matrices for CEI-II in Study 1: Full set of variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Stretching 1	-.215	.748	.230	.621	.415
Stretching 2	.401	.194	.516	.432	.291
Stretching 3	.274	.494	.567	.656	.479
Stretching 4	.698	-.047	.669	.367	.450
Stretching 5	.431	.363	.646	.619	.503
Embracing 1	.018	.049	.047	.060	.004

Embracing 2	.387	.330	.583	.559	.410
Embracing 3	.807	-.164	.710	.315	.521
Embracing 4	.809	-.121	.737	.360	.553
Embracing 5	-.064	.380	.162	.343	.120
					Total
Sum of Squares (eigenvalues)	3.102	.644			3.746
Explained Variance	31.017	6.441			37.458

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.
* Factor loadings more than .3 have been given in bold

Table 18

Promax-Rotated Pattern and Structure Matrices for CEI-II in Study 1: Reduced set of variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Stretching 1	-.187	.590	.132	.488	.263
Stretching 3	.068	.766	.483	.803	.649
Stretching 5	.248	.456	.495	.590	.392
Embracing 2	.311	.473	.567	.641	.480
Embracing 3	1.035	-.206	.923	.354	.883
Embracing 4	.514	.179	.611	.457	.396
					Total
Sum of Squares (eigenvalues)	2.138	.923			3.061
Explained Variance	35.632	15.388			51.020

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.
* Factor loadings more than .3 have been given in bold

After Promax rotation, although the same items loaded on factors strongly, it was observed that Embracing 2 cross loaded on two factors and when it was deleted, the structure of the model was collapsed. Another problem in this analysis was about the high factor loading Embracing 3 had. As a result, it was assumed that using Embracing sub section of CEI-II might not be appropriate as suggested by authors.

When the validity and reliability of the scale were examined, it could be realized that all items in Factor 2 (Stretching) highly loaded only in this factor. Therefore, it would not be wrong to state that this factor was unidimentional. Its reliability, its Cronbach-alpha value was .66, which indicated the existence of problems. Since the problems with Factor 1 (Embracing) were not solved, there was no need to check the factor correlations for discriminant validity. All in all, when the number of cases, which was 92, was taken into consideration, this might have resulted from relatively low number of participants. Thus, it could be said that CEI-II should be improved before using it as a scale as suggested by its authors.

EFA with Achievement Goal Questionnaire – Revised. In the final EFA analysis of Study 1, factor analysis for 'Achievement Goal Questionnaire – Revised' which had two subdimensions (Mastery Avoidance and Performance Avoidance) to be used in the scope of this study was carried out. In terms of the suitability of the data set for factor analysis, Pearson correlation matrix was examined. It was found that most of the correlation coefficients were above .3. Of the 23 correlations, 10 were significant at .01 (2-tailed) level while 1 of them was significant at .05 (2-tailed) level. This finding showed the existence of many significant correlations which would worth investigating further. In addition, Bartlett's test of sphericity and chi-square = 193.071, df = 15, $p < .001$ indicated the significance of the correlations. Regarding measure of sampling adequacy, Kaiser-Meyer-Olkin measure of sampling adequacy value was found to be .691. This was a value above the suggested limit by Kaiser (1974) and was accepted as average by Hair et al. (2013). These findings indicated that the data were suitable for factor analysis.

Maximum likelihood extraction method provided two factors with eigenvalues above 1 as expected. Table 19 gives eigenvalues, percentages of explained variance and unrotated factor loadings for each factor. As indicated, Factor 1 had 2.912 eigenvalue and explained the 48.529% of the total variance without any rotation method was applied. On the other hand, Factor 2 explained 19.865% of the variance with 1.192 eigenvalue. Factor 1 and Factor 2 explained the 68.395% of the total variance with a total 4.104 eigenvalue. When the communalities were checked, only one item had a communality value below .4, which is out of the limit suggested by Osborne, Costello and Kellow (2008). However, in order not to weaken the structure and lose this section of the survey, no action was taken at this stage and this item was followed carefully throughout the factor analysis.

Table 19

Unrotated Factor Matrix for AGQ-R in Study 1

	Factor		Communality
	1	2	
Mastery Avoidance 1	.999	-.002	.999
Mastery Avoidance 2	.669	.186	.483
Mastery Avoidance 3	.279	.293	.164
Performance Avoidance 1	.466	.581	.555
Performance Avoidance 2	.249	.784	.676
Performance Avoidance 3	.176	.770	.624
			Total
Sum of Squares (eigenvalues)	2.912	1.192	4.104

Explained Variance	48.529	19.865	68.395

Extraction Method: Maximum Likelihood.

Although Table 19 provides a clear picture in terms of the factor loadings, there were some problems with items Mastery Avoidance 3 and Performance Avoidance 1 which cross loaded onto the two factors. Therefore, promax rotation was carried out. After the rotation, the communality values for each variable were checked and it was seen that the communality value of Mastery Avoidance 3 was still below the suggested level. Table 20 shows the promax-rotated pattern and structure matrixes for full and reduced set of variables together with eigenvalues of the factors and their percentages of explained variance after the rotation.

Table 20

Promax-Rotated Pattern and Structure Matrix for AGQ-R in Study 1: Full set of variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Mastery Avoidance 1	-.098	1.036	.327	.995	.999
Mastery Avoidance 2	.139	.626	.396	.683	.483
Mastery Avoidance 3	.293	.183	.369	.303	.164
Performance Avoidance 1	.590	.273	.702	.515	.555
Performance Avoidance 2	.832	-.025	.822	.317	.676
Performance Avoidance 3	.825	-.096	.785	.243	.624
					Total
Sum of Squares (eigenvalues)	1.835	1.665			3.500
Explained Variance	30.581	27.756			58.337

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

As can be seen in Table 20, Mastery Avoidance 3 item had low factor loading which was below the suggested levels and loaded onto the Factor 1 although it was expected to load on Factor 1. Thus, the item (Mastery Avoidance 3) was eliminated and model was respecified. Table 21 gives the figures for promax-rotated pattern and structure matrices, communality values, eigenvalues and percentage of the explained variance for each factor for the reduced set of variables. After the elimination of Master Avoidance 3, communality values were improved and all were above .5. In addition, the eigenvalue for Factor 1 increased to 2.377 and it explained 47.545% of the total variance whereas the eigenvalue for Factor 2 dropped to .872 and the percentage of the explained variance by this factor was 17.444%.

Table 21

Promax-Rotated Pattern and Structure Matrices for AGQ-R in Study 1: Reduced set of variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Mastery Avoidance 1	-.077	.895	.329	.860	.745
Mastery Avoidance 2	.034	.764	.381	.780	.609
Performance Avoidance 1	.584	.304	.721	.569	.594
Performance Avoidance 2	.761	.028	.773	.373	.598
Performance Avoidance 3	.897	-.154	.827	.253	.703
					Total
Sum of Squares (eigenvalues)	2.377	.872			3.249
Explained Variance	47.545	17.444			64.990

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

When Table 21 is examined, it is seen that Performance Avoidance 1 is loading on both factors. It loaded onto both on Factor 1 (.584) and Factor 2 (.304) with a difference of .280. It was evident that this item was strongly loaded onto Factor 1. Thus, it was not necessary to respecify the model.

Since all the variables loaded strongly onto one factor (providing Performance Avoidance 1 belongs to Factor 1), unidimentionality was achieved. In terms of reliability, It was found that Factor 1 (Performance Avoidance) had a reliability value of Cronbach-alpha = .804 whereas the Cronbach-alpha value for Factor 2 (Mastery Avoidance) was .801. Accordingly, it can be concluded that this is a reliable scale (Hair et al., 2013). Regarding the validity of the scale, first its convergent validity was checked. As can be seen in Table 21, all the variables were loaded strongly, which meant that convergent validity was achieved. For discriminant validity, factor correlations, which were given in Table 22, were investigated.

Table 22

Promax-Rotated Factor Correlation Matrix for AGQ-R in Study 1

Factor	1	2
1	1.000	.454
2	.454	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

According to Table 22, the two factors had a correlation of .454, which was in the acceptable range that signals the existence of discriminant validity.

In this respect, it could be said that the scale could be used as separate scales as suggested by the authors. But since the number of items in Mastery Avoidance decreased, it would be underidentified for SEM analysis.

Study 2

Data Screening

Type of Missing Data. Like Study 1, before proceeding for further analyses, first missing data were examined. As a result, it was figured out the reasons for the missing data were unknown. That is, they were not due to wrong entries during data entry stage but because participants had chosen not to respond to those items. In the next stage, the extent of missing data was investigated to see whether it was above the suggested limits (Hair et al., 2013).

Determination of the extent of missing data. At this stage, the percentage of the missing data per case and variable was investigated. As in Study 1, demographic data were excluded since they were not going to be used for further analyses. Table 23 shows the distribution of missing data per variable in a descending order. Examination of missing values per variable revealed that out of 117 variables 23 (19.66%) had at least one missing value and the total number was 28. Before applying any imputation method, missing values per case were also investigated.

Table 23

Summary of Missing Data per Variable in Study 2

	N	Mean	Std. Deviation	Missing	
				Count	Percent
Belonging 4	335	2.99	1.096	3	.9
Mastery Avoidance 1	336	3.84	.946	2	.6
Belonging 1	336	2.93	1.183	2	.6
Control 1	336	3.42	1.031	2	.6
TDMC5	337	2.81	1.553	1	.3
Mastery Avoidance 2	337	3.91	.888	1	.3
Autonomy Frustration 1	337	3.33	1.183	1	.3
Relatedness Frustration 1	337	1.80	.934	1	.3
Competence Satisfaction 2	337	3.61	.926	1	.3
Autonomy Frustration 3	337	2.60	1.221	1	.3
Relatedness Frustration 3	337	2.05	1.161	1	.3
Competence Satisfaction 3	337	3.88	.868	1	.3
Competence Frustration 3	337	2.51	1.225	1	.3
Competence Satisfaction 4	337	3.86	.884	1	.3
Embracing	337	2.29	1.260	1	.3

Scarcity 1	337	4.11	1.981	1	.3
GLB Entity 1	337	2.73	1.442	1	.3
GLB Entity 2	337	2.60	1.434	1	.3
GLB Entity 3	337	2.29	1.414	1	.3
L2B Incremental 1	337	4.40	1.306	1	.3
Psychological Ownership 4	337	2.64	1.009	1	.3
Belonging 2	337	2.56	1.166	1	.3
Proud 4	337	3.30	1.263	1	.3

Table 23 shows the distribution of missing values for each case. Accordingly, it could be seen that out of 338 cases, 20 cases had missing values. Two of them had three missing values, four of them had two missing values and 14 of them had only one missing value.

Table 24

Amount of missing data per case in Study 2

Number of Missing Data per Case	Number of Cases	Percent
0	318	94.08
1	14	4.14
2	4	1.18
3	2	0.59
Total	338	100

When the number of items per case, which was 117, was taken into consideration, the number of missing values per case was not problematic. That is, the number of missing values was below the suggested level (Hair et al., 2013). Thus, in the next step, the randomness of the missing values was tested.

Randomness of the missing data. According to the findings, the number of missing values per case and variable was low enough to proceed. However, as pointed out by Hair et al. (2013) apart from diagnosing the amount of missing values, it is also necessary to examine the patterns of the data empirically to choose correct imputation methods. Little's MCAR test, chi-square = 2846.198, df = 2426, sig. = .000, indicated that missing data were not distributed in a random pattern (MAR).

Selecting an imputation method. Hair et al. (2013) argue that when missing data is lower than 10%, any kind of imputation method can be applied (p. 54). As an imputation method, EM approach, which is an iterative two-stage method (Hair et al., 2103, p. 48), was applied and mean values obtained from EM statistics were used for the missing data.

Outliers. As in Study 1, at this stage first the data set was scanned for entry errors. A frequency analysis was done and as a result, entries with extreme values and the cases that included those values were identified. Then, for each problematic value, original entry was checked by using the ID codes of the cases given during data entry stage. In this way, it was possible to fix the wrong entries.

At the next stage, cases with unengaged responses were identified. For this purpose, SDs for each case were calculated. It was assumed that if a participant had made the same choice for all the items, the standard deviation for that case must have been equal to zero. It was found out that the lowest SD value was 0.7434754, which meant that all the participants replied to the items in the questionnaire in an engaged way.

Regarding other types of outliers such as univariate, bivariate and multivariate ones, none of the outlier detection methods were applied because, first, when the formation of the items in the data set was taken into consideration, all of them were representatives of observations and second, there were no continues variable in them. And finally, as Hair et al. (2013) argued, all the outliers were retained to ensure the generalizability to the entire population.

Statistical Assumptions

Normality. It is known that for all multivariate data analyses, the data should be normally distributed across variables. Basically, there are two types of normality test: univariate and multivariate. When individual variables are tested, it is called univariate normality whereas when two or more variables are tested together, it is called multivariate normality. As Hair et al. (2013) indicate, normal distribution of the data is closely related to the size and shape of the distribution. The larger the sample size is, the more possible it is to overcome the negative consequences of the normality problems. In this vein, when the sample size of the data in Study 2 (338) was taken into consideration, it was possible to state that the sample size was large enough.

Regarding the shape of the distribution of the data, the Skewness and Kurtosis values, which can be used to diagnose the normal distribution, should be checked. It is assumed that when these values are equal to zero, it is probable to accept that the data are distributed normally across variables. As addressed in

Study 1, there are different points of views in terms of what the acceptable Skewness and Kurtosis values should be. Çokluk et al. (2012) state that Skewness and Kurtosis values between ±1 are enough to assume the existence of normality. On the other hand, according to Hair et al. (2013) at 0.5 significance level, values ±1.96 indicate normal distribution. There are researchers who adopt looser limits. George and Mallery (2013) and Civelek (2017) point out that for normality, Skewness and Kurtosis values of ±2 are sufficient. Sposito et al., (1983) state it should be between ±2.2 whereas Kline (2011) states values between ±10 can be accepted. Table 25 shows the Skewness and Kurtosis values for the 50 variables used in Study 2.

Table 25

Skewness and Kurtosis Values of Study 2

	N		Skewness	Std. Error of Skewness	Kurtosis	Std. Error of Kurtosis
	Valid	Missing				
Scarcity 5	338	0	-1.598	.133	2.157	.265
Scarcity 2	338	0	-1.327	.133	.825	.265
Scarcity 3	338	0	-1.254	.133	.703	.265
Mastery Avoidance 2	338	0	-1.149	.133	1.820	.265
Relatedness Satisfaction 4	338	0	-.992	.133	.896	.265
Relatedness Satisfaction 3	338	0	-.979	.133	1.184	.265
Mastery Avoidance 1	338	0	-.949	.133	.872	.265
DMC1	338	0	-.896	.133	-.511	.265
Mastery Avoidance 3	338	0	-.895	.133	1.244	.265
DMC11	338	0	-.865	.133	-.621	.265
DMC2	338	0	-.858	.133	-.577	.265
Relatedness Satisfaction 2	338	0	-.841	.133	.262	.265
Performance Avoidance 2	338	0	-.822	.133	.120	.265
Competence Satisfaction 1	338	0	-.774	.133	.681	.265
DMC10	338	0	-.763	.133	-.739	.265
Performance Avoidance 1	338	0	-.731	.133	-.258	.265
Competence Satisfaction 4	338	0	-.715	.133	.692	.265
Autonomy Satisfaction 3	338	0	-.705	.133	.742	.265
DMC5	338	0	-.699	.133	-.542	.265
Autonomy Satisfaction 1	338	0	-.632	.133	-.103	.265
Competence Satisfaction 2	338	0	-.605	.133	.505	.265
DMC4	338	0	-.585	.133	-.999	.265
Performance Avoidance 3	338	0	-.581	.133	-.421	.265
Stretching 3	338	0	-.559	.133	-.058	.265
Competence Satisfaction 3	338	0	-.538	.133	.252	.265
Embracing 2	338	0	-.524	.133	-.111	.265
Relatedness Satisfaction 1	338	0	-.514	.133	.000	.265
DMC12	338	0	-.442	.133	-.984	.265
DMC9	338	0	-.435	.133	-.877	.265
Autonomy Satisfaction 4	338	0	-.432	.133	-.330	.265
DMC3	338	0	-.404	.133	-.885	.265
DMC6	338	0	-.392	.133	-.820	.265
Autonomy Satisfaction 2	338	0	-.358	.133	.272	.265
Stretching 5	338	0	-.303	.133	-.527	.265

Stretching 1	338	0	-.275	.133	-.128	.265
Embracing 5	338	0	-.241	.133	-.763	.265
DMC7	338	0	-.237	.133	-1.082	.265
Embracing 3	338	0	-.211	.133	-.801	.265
Psychological Ownership 3	338	0	-.187	.133	-.763	.265
Stretching 2	338	0	-.176	.133	-.017	.265
Psychological Ownership 2	338	0	-.161	.133	-.403	.265
Psychological Ownership 1	338	0	-.155	.133	-.680	.265
Stretching 4	338	0	-.138	.133	-.593	.265
DMC8	338	0	-.127	.133	-1.093	.265
Embracing 4	338	0	-.102	.133	-.661	.265
Scarcity 1	338	0	-.034	.133	-1.149	.265
Psychological Ownership 5	338	0	-.003	.133	-.722	.265
Psychological Ownership 4	338	0	.061	.133	-.404	.265
Scarcity 4	338	0	.377	.133	-1.030	.265
Embracing 2	338	0	.644	.133	-.646	.265

As can be seen in Table 25, regarding the Skewness values, there were only four values outside ±1 range whereas there were eight Kurtosis values outside the same range. And none of the values were above 2 or below -2. Since none of these values were extreme ones, no action was taken and it was decided to retain all the variables. Based on these findings, it could be stated that Skewness (highest .644, lowest -1.598) and Kurtosis (highest 2.157, lowest -1.149) values were between the published thresholds and it was possible to assume that univariate normality was achieved. In other words, normality assumption was not violated.

Linearity. As in Study 1, it was possible to accept the existence of curvilinearity between some variables due to the differences between Skewness values. Hair et. al, (2013) and Tabachnick and Fidell (2013) suggest checking scatterplots to diagnose such kind of issues. However, with 50 variables which result in over 1000 scatterplots it was not practical to check all the plots (Tabachnick & Fidell, 2013; p. 657). Thus, only some variables with highest and lowest Skewness values were checked.

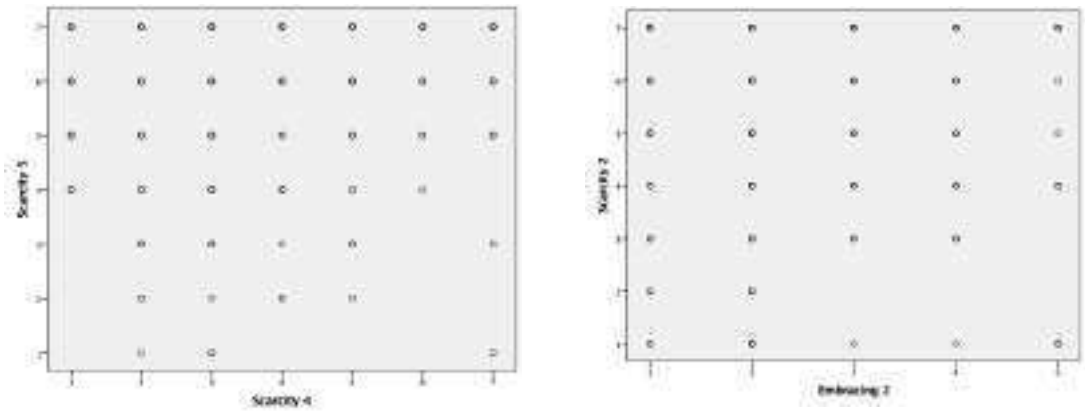


Figure 11. Scatterplots for linearity check among variables in Study 2.

Figure 11 shows sample scatterplots which have the highest negative (Scarcity 5 and Scarcity 2) values and the highest positive (Scarcity 4 and Embracing 2) values. Due to their Skewness values, these variables were expected among the worst. However, when the scatterplots were investigated, it was not possible to say that linearity existed between these variables. But it was also not possible to say that there was not enough evidence that show curvilinearity. Therefore, no action was taken in terms of data transformation.

Exploratory Factor Analysis in Study 2.

In Study 1, EFA analyses resulted in a reduced version of the chosen questionnaires and scales. Consequently, it was found out that scales and questionnaires could be utilized as suggested by their authors in that reduced version except the Curiosity and Exploration Inventory (Kashdan et al., 2009). That was one of the reasons that led to the administration of Study 2. In addition, although it was possible to use subscales Scarcity, Competence Satisfaction and Mastery Avoidance as suggested, the number of items were not enough to use those subscales in a further SEM analyses since they would become underidentified. However, what underlies these problems might be the relatively small sample size (n = 92). Thus, it would be better to pilot the adapted scales and questionnaires once more. In addition, another reason that rationalizes Study 2 is the feedback obtained from both participants and consulted experts. Because those feedbacks were pointing out severe face validity issues which might have interfered with the results. In accordance with participants' and experts' feedbacks, the paper size, font style and size, numbers of the items, and section titles were changed. Moreover, adjustments regarding paper layout had been made so that participants would find completing the questionnaire easy. Thus, it was decided to pilot the questionnaire in a revised version.

Assumptions and procedures in factor analysis. As a result of the above-mentioned data screening procedure, there were 338 cases with 117 variables. Before starting exploratory factor analyses, first, data set was trimmed and the variables which would not be used were deleted. In addition, cases in which participants stated they had never experienced any kind of DMC were also excluded. Finally, a new data set with 278 cases and 50 variables all of which

represented eight core drives of Octalysis framework and were grouped under ten subscales was formed. When the suggested threshold (Hair et al., 2013) which was 100 and the 5 to 1 ratio were taken into consideration, it could be said that the sample size in Study 2 was enough to conduct factor analyses.

As in the previous stage, different exploratory factor analysis procedures were applied for each subscale with two or more dimensions. In this respect, first EFA included DMC, Scarcity and Psychological Ownership dimensions. These were different from each other and each had only one dimension. Then, in the subsequent analyses, items that represent satisfaction dimensions of BPNSF-General, CEI-II and AGQ-R were tested separately. The assumptions those had to be met before factor analyzing such as correlations among variables, significance level of correlations, and sampling adequacy were all tested and reported before each EFA procedure. In all EFAs, Maximum Likelihood Extraction (MLE) method (Lawley & Maxwell, 1963) was used. Then, Promax Kaiser Rotation was applied and consequently findings were evaluated.

EFA with DMC, Scarcity and Psychological Ownership. Before starting factor analysis, assumptions were checked to see whether the data were suitable for factor analysis. To be able to conduct a factor analysis, variables should be correlated with each other. To test this assumption, a correlation matrix was created and it was found that there were correlations of r = .3 or above (Pallant, 2011). In addition, when the matrix (Appendix 5) was examined, it could be seen that 67 correlations were significant at .000 level and 23 correlations were significant at .05 level. Correlation matrix showed that there were enough significant correlations among variables to conduct a factor analysis. Another assumption to be met before proceeding was to test the significance of these correlations by checking the Bartlett's Test of Sphericity (Bartlett, 1954). Bartlett's test, chi-square = 1691,020, df = 231, p < .000, indicated that variables correlated significantly. The third assumption was testing whether the sampling was adequate or not. The Kaiser-Meyer-Olkin measure of sampling adequacy value was .797, which was adequate (Hair et al., 2013) to go on with the factor analysis.

Factor analysis was carried out by using maximum likelihood extraction method. However, instead of determining the number of factors according to the eigenvalues, a priori criterion (Hair et al., 2013; p. 107), an approach which could

be used while testing theories, was adopted and factor extraction stopped when three factors were obtained. Table 26 presents the unrotated factor matrix for DMC, Scarcity and Psychological Ownership in Study 2.

Table 26

Unrotated Factor Matrix for DMC, Scarcity and Psychological Ownership in Study 2

	Factor			Communality
	1	2	3	
DMC1	.482	.346	-.141	.372
DMC2	.428	.391	.206	.378
DMC3	.372	.311	.208	.278
DMC4	-.146	.006	.231	.075
DMC5	.480	.395	-.097	.395
DMC6	.549	.431	-.154	.511
DMC7	.364	.393	-.104	.298
DMC8	.415	.376	-.011	.314
DMC9	.384	.339	-.042	.264
DMC10	.388	.372	.014	.289
DMC11	.347	.310	.057	.220
DMC12	.510	.250	-.156	.347
Scarcity 1	.255	-.124	.300	.170
Scarcity 2	.117	.019	.378	.157
Scarcity 3	.158	-.002	.645	.440
Scarcity 4	.251	.116	.008	.077
Scarcity 5	.174	.144	.509	.311
Psychological Ownership 1	.707	-.461	.025	.713
Psychological Ownership 2	.621	-.368	-.063	.525
Psychological Ownership 3	-.518	.316	.030	.369
Psychological Ownership 4	.286	-.292	-.107	.178
Psychological Ownership 5	.647	-.427	-.036	.602
				Total
Sum of squares (eigenvalues)	4.761	2.512	1.840	9.113
Explained Variance	21.642	11.418	8.362	41.421

Extraction Method: Maximum Likelihood.

When Table 26 is investigated, it is seen that the three factors had eigenvalues of 4.761, 2.512, and 1.840, respectively and they explained 41.421% of the total variance without any rotation. However, the unrotated factor matrix was helpful to obtain clear factor structures, which meant that rotation was necessary. But, before proceeding, communalities were examined and there were many communalities below the suggested thresholds. Hair et al. (2013) suggest that communalities should not drop below .5 level. On the other hand, Osborne et al. (2008) argue that .4 could be used as a cut-off point whereas Child (2006) states that communalities below .2 should be eliminated. Therefore, before proceeding with the factor analysis with rotation, items with low level of communality values were deleted. In the first run, DMC 4, Scarcity 4 and Psychological Ownership 4

were deleted and analysis was rerun. Consequently, it was observed that unrotated eigenvalues and the total variance explained by the three factors increased considerably. After deletion of three items, the factors explained 46.165% of the total variance. Although there were still some items (Scarcity 1 and Scarcity 2) with communalities below .2, no action was taken to delete those two factors since it would have caused to lose whole Scarcity sub-section. Thus, factor analysis was not stopped.

In the next step, Promax Kaiser Normalization Rotation was applied to get clearer set of factor loadings. After rotation, it was seen that all the items significantly loaded on a factor with a loading higher than .3. Promax rotation yielded a clear set of structures and there was no need to apply any more procedures. Table 27 and 28 show the pattern and structure matrices, communalities, eigenvalues and percentage of explained total variance for the three factors after the Promax rotation. From the Tables 27 and 28, it could be seen that after rotation, the total variance that the three factors explained increased to 36.786 from 33.111.

Table 27

Promax-Rotated Pattern and Structure Matrices for DMC, Scarcity and Psychological Ownership in Study 2: Full set of variables

	Pattern Matrix*			Structure Matrix*			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
DMC1	.604	.052	-.102	.600	.220	.022	.372
DMC2	.535	-.064	.247	.564	.126	.342	.378
DMC3	.433	-.035	.244	.470	.124	.324	.278
DMC4	-.129	-.129	.223	-.124	-.140	.182	.075
DMC5	.636	.006	-.057	.626	.188	.067	.395
DMC6	.719	.029	-.109	.707	.229	.034	.511
DMC7	.571	-.067	-.074	.536	.093	.028	.298
DMC8	.564	-.032	.024	.559	.139	.130	.314
DMC9	.521	-.018	-.010	.514	.136	.089	.264
DMC10	.540	-.050	.047	.534	.117	.146	.289
DMC11	.453	-.032	.088	.460	.114	.172	.220
DMC12	.540	.150	-.115	.562	.297	.008	.347
Scarcity 1	-.032	.229	.326	.099	.261	.349	.170
Scarcity 2	-.003	.013	.395	.077	.061	.396	.157
Scarcity 3	-.058	.023	.670	.078	.089	.661	.440
Scarcity 4	.240	.069	.030	.267	.144	.085	.077
Scarcity 5	.109	-.068	.534	.192	.032	.547	.311
Psychological Ownership 1	-.012	.833	.083	.252	.840	.185	.713
Psychological Ownership 2	.040	.713	-.014	.250	.723	.084	.525
Psychological Ownership 3	.022	.599	.011	.202	.607	.090	.369
Psychological Ownership 4	-.071	.438	-.086	.043	.406	-.045	.178
Psychological Ownership 5	-.002	.774	.016	.232	.776	.112	.602
							Total

Sum of squares (eigenvalues	3.941	2.185	1.158		7.284
Explained Variance	17.914	9.931	5.265		33.111

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.
*Factor loading above .3 have been given in bold

Table 28

Promax-Rotated Pattern and Structure Matrices for DMC, Scarcity and Psychological Ownership in Study 2: Reduced set of Variables

	Pattern Matrix*			Structure Matrix*			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
DMC1	.593	.066	-.072	.597	.229	.069	.364
DMC2	.522	-.059	.263	.563	.136	.369	.382
DMC3	.427	-.018	.225	.472	.143	.317	.270
DMC5	.623	.023	-.043	.620	.199	.098	.387
DMC6	.702	.037	-.079	.696	.231	.082	.491
DMC7	.577	-.045	-.114	.538	.106	.006	.305
DMC8	.570	-.016	-.007	.564	.151	.116	.318
DMC9	.527	-.020	-.015	.518	.132	.098	.269
DMC10	.557	-.052	.029	.548	.117	.144	.304
DMC11	.462	-.030	.078	.471	.118	.175	.228
DMC12	.529	.138	-.065	.556	.283	.073	.329
Scarcity 1	-.054	.240	.329	.090	.276	.355	.178
Scarcity 2	-.028	.012	.419	.069	.070	.415	.173
Scarcity 3	-.089	.016	.700	.070	.099	.683	.473
Scarcity 5	.091	-.063	.528	.189	.047	.538	.298
Psychological Ownership 1	-.010	.820	.059	.244	.826	.185	.685
Psychological Ownership 2	.030	.735	-.042	.237	.738	.080	.547
Psychological Ownership 3	.002	.631	-.005	.187	.631	.095	.399
Psychological Ownership 5	-.005	.768	.009	.223	.768	.128	.590
							Total
Sum of squares (eigenvalues	3.831	2.049	1.110				6.990
Explained Variance	20.162	10.782	5.842				36.786

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.
* Factor loadings more than .3 have been given in bold

In addition, according to Tables 27 and 28, it is understood that all variables loaded highly on a factor, which meant unidimensionality was obtained because of the Promax rotation. In other words, there were strong associations between items and it was possible to get a single dimension (Hair et al., 2013; p. 123).

Having found out that Factors 1, 2 and 3 were unidimensional, further analyses were carried out to see how reliable and valid they were. For this purpose, the Cronbach's alpha values were checked. Cronbach's alpha for DMC was .830. For Scarcity, it was .545 and finally it was .827 for Psychological Ownership. Hair et al. (2013) argue that the Cronbach's alpha should be higher than .7. In this respect, it could be stated that Scarcity subscale had some issues regarding its reliability.

In the next step, validities of these three factors were checked. As in Study 1, issues related to face validity were dealt with at the end of Study 2 after all separate EFAs were carried out. In terms of convergent validity, item loadings should be high for a single factor. When Table 28 is controlled, it is seen that all the variables highly loaded on Factors. That is, convergent validity was obtained for all factors. The other validity check was discriminant validity. According to Gaskin (2018) and Farrell and Rudd (2009), for discriminant validity, factors need to be uncorrelated and distinct. Table 29 presents Promax-Rotated factor correlation matrix.

Table 29

Promax-Rotated Factor Correlation Matrix for DMC, Scarcity and Psychological Ownership in Study 2

Factor	1	2	3
1	1.000	.294	.221
2	.294	1.000	.156
3	.221	.156	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

As can be understood from Table 29, there were low correlations among factors, which indicates the existence of different factors. Thus, it was assumed that these factors had discriminant validity.

EFA with Basic Psychological Needs Satisfaction and Frustration Scale. Second EFA procedure in Study 2 was carried out by testing the subscales of BPNFS-General. In accordance with the aim of this study, only the subscales Competence Satisfaction, Relatedness Satisfaction and Autonomy satisfaction were included into the factor analysis since hypothetically these three subscales were assumed to be representatives of the core drives of Octalysis which are Development and Accomplishment (CD 2), Social İnfluence and Relatedness (CD 5), and Empowerment of Creativity and Feedback (CD 3). First, the appropriateness of the data set for factor analysis was tested. In this vein, the correlation matrix (Appendix 6) was examined. Findings revealed the existence of many correlations of r = .3 or above. Out of 78 correlations, 46 were significant at .01 level (2-tailed) and seven were significant at .05 level (2-tailed). Bartlett's Test of Sphericity, approx. chi-square = 722.313, df = 66, $p < .000$, also indicated that correlations were significant enough to conduct factor analysis. Kaiser-Meyer-Olkin measure of

sampling adequacy was found .835, which was accepted as meritorious by Hair et al. (2013). When these findings were taken into consideration, it was assumed that the data was suitable for factor analysis.

In the next step, factor analysis with maximum likelihood extraction method was conducted. As in the previous section and Study 1, a priori criterion was applied and factor extraction was stopped when three factors were obtained. Table 30 gives the unrotated factor loadings, communalities of each variable, eigenvalues and percentages of explained total variance for each factor.

Table 30

Unrotated Factor Matrix for BPNSF-General in Study 2

	Factor			Communality
	1	2	3	
Autonomy Satisfaction 1	.249	.459	-.035	.274
Autonomy Satisfaction 2	.153	.588	.195	.408
Autonomy Satisfaction 3	.250	.658	.263	.564
Autonomy Satisfaction 4	.149	.621	-.029	.409
Relatedness Satisfaction 1	.313	.320	.406	.365
Relatedness Satisfaction 2	.215	.111	.342	.176
Relatedness Satisfaction 3	.999	-.002	.000	.999
Relatedness Satisfaction 4	.356	.185	.176	.192
Competence Satisfaction 1	.005	.521	-.204	.313
Competence Satisfaction 2	.135	.581	-.109	.368
Competence Satisfaction 3	.156	.572	-.291	.436
Competence Satisfaction 4	.144	.507	-.335	.390
				Total
Sum of Squares (eigenvalues)	3.764	1.611	.952	6.327
Explained variance	31.369	13.429	7.937	52.735

Extraction Method: Maximum Likelihood.

Table 30 shows that it was not possible to get clear factor structures without applying rotation. Although it could be seen that there were some variables highly loaded on a single factor, there were also some problems. For example, it was not possible to reach a conclusion in terms of on which factor Relatedness Satisfaction 1 and 2, Autonomy Satisfaction 1 and 3, and Competence Satisfaction 4 were loaded. In addition, when the communalities were examined, it was seen that Relatedness Satisfaction 2 and 4 had low communalities. Another problem that could be understood from Table 30 was about the eigenvalues that each factor had. As a generally accepted rule, eigenvalues should be higher than 1. But in this case eigenvalue for factor 3 was below that threshold. Eigenvalues obtained for each factor were: 3.764, 1.611and .952, respectively. The three factors were also explaining 52.735% of the total variance without any rotation. Consequently, before

moving on with further analysis, first the item with lowest communality (Relatedness Satisfaction 2) was eliminated. In order not to produce underidentified group of items for further analyses, the other item with a communality value below the suggested threshold was not deleted. After all, to be able to get the optimum factor structures Promax rotation was applied.

After Promax rotation, small changes were seen in terms of total variance explained by the three factors. The eigenvalue for Factor 3 dropped to .553 and total variance explained was 42.920%. When factor loadings were examined, it was seen that Autonomy satisfaction was cross-loaded on Factors 1 and 2. The problematic variables were eliminated until a clear set of structures were obtained. As a result, it was found out that subscale Relatedness Satisfaction could not be used as suggested by the authors. Thus, this subscale was completely excluded from the analysis. Tables 31 and 32 present the pattern and structure matrices together with the sum of square values and percentages of explained variance by the two remaining factors.

Table 31

Promax-Rotated Pattern and Structure Matrices for BPNSF in Study 2: Full set of variables

	Pattern Matrix*			Structure Matrix*			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
Autonomy Satisfaction 1	.412	.122	.108	.496	.390	.201	.274
Autonomy Satisfaction 2	.304	.434	-.097	.539	.578	.077	.408
Autonomy Satisfaction 3	.302	.549	-.046	.609	.707	.163	.564
Autonomy Satisfaction 4	.535	.168	-.040	.626	.461	.084	.409
Relatedness Satisfaction 1	-.100	.629	.072	.268	.594	.252	.365
Relatedness Satisfaction 2	-.214	.477	.063	.066	.374	.180	.176
Relatedness Satisfaction 3	.029	.107	.955	.219	.417	.992	.999
Relatedness Satisfaction 4	.001	.311	.227	.208	.381	.322	.192
Competence Satisfaction 1	.608	-.093	-.091	.544	.226	-.037	.313
Competence Satisfaction 2	.574	.057	-.015	.605	.380	.080	.368
Competence Satisfaction 3	.734	-.168	.065	.647	.270	.112	.436
Competence Satisfaction 4	.720	-.243	.088	.594	.195	.110	.390
							Total
Sum of Squares (eigenvalues)	1.502	2.718	.674				4.894
Explained variance	12.521	22.646	5.614				40.781

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Loadings higher than .3 were given in bold

Table 32

Promax-Rotated Pattern and Structure Matrices for BPNSF-G in Study 2: Reduced set of variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Competence Satisfaction 1	.208	.363	.418	.484	.263
Competence Satisfaction 3	.180	.537	.491	.641	.433
Competence Satisfaction 4	-.104	.812	.365	.752	.572
Autonomy Satisfaction 1	.387	.137	.466	.361	.230
Autonomy Satisfaction 2	.634	.035	.654	.401	.429
Autonomy Satisfaction 3	.798	-.057	.765	.404	.587
					Total
Sum of Squares (eigenvalues)	2.065	.449			2.514
Explained variance	34.419	7.487			41.906

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

According to Table 32, it was possible to obtain a clear set of factor structures after eliminating the relatedness dimension of the original scale. Autonomy Satisfaction 1, 2 and 3 formed Factor 1 whereas Competence Satisfaction 1, 3 and 4 formed the second factor. Since these items loaded significantly on a single factor, it could be argued that these two factors were unidimensional. Reliability check for these factors revealed that Factor 1 (Autonomy Satisfaction), Cronbach's Alpha = .641; and Factor 2 (Competence Satisfaction), Cronbach's Alpha = .651 had reliability values below the satisfactory level. However, Hair et al. (2013) state that in exploratory studies, it is possible to decrease the cut-off point to .6 to determine the reliability.

In terms of the validity checks for these variables, item loadings in Factors 1 and 2 were high enough to assume that convergent validity was obtained. Regarding discriminant validity, factor correlations given in Table 33 were controlled.

High factor correlations provided in Table 33 indicate problems in terms of discriminant validity because if there are high correlations (.578 in this case) between factors, it means that they are basically not different from each other.

Table 33

Factor Correlation Matrix for BPNSF-G in Study 2

Factor	1	2
1	1.000	.578
2	.578	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

Therefore, as a result of these findings, it can be concluded that although it was possible to get clear set of factors with high loadings, for further analyses the validity and reliability issues may cause problems while using Competence Satisfaction and Autonomy Satisfaction subscales as suggested by their authors.

EFA with The Curiosity and Exploration Inventory 2. First, the data were investigated regarding the statistical assumptions. In this respect, when the correlation matrix (Appendix 7) was investigated, it was seen that there were many correlations of .3 or above and of the 55 correlations 34 were significant at .01 level (2-taied) and four were significant at .05 level (2-tailed). In addition, Bartlett's Test of Sphericity, approx. chi-square = 625.890, df = 45, $p < .001$; and Kaiser-Meyer-Olkin measure of sampling adequacy value which was .847 indicated that it was appropriate to apply factor analysis.

Maximum likelihood extraction method was applied and since the number of factors to be expected was known, the extraction was stopped when 2 factors were obtained. Before proceeding, unrotated factor loadings, communality values and eigenvalues were examined. Accordingly, it was found out that there were some problems with the communalities. Communality values for Stretching 1 (.158) and Embracing 1 (.053) were especially below the suggested level (Child, 2006; Osborne et al., 2008). Besides, it was seen that Factor 1 had 3.603 eigenvalue while Factor 2 had 1.184. Initially, the two factors explained 47.866% of the total variance.

Table 34

Unrotated Factor Matrix for CEI-II in Study 2

	Factor		Communality
	1	2	
Stretching 1	.196	.347	.158
Stretching 2	.504	.194	.292
Stretching 3	.617	.330	.489
Stretching 4	.680	-.201	.503
Stretching 5	.622	.023	.388
Embracing 1	.220	-.068	.053
Embracing 2	.594	.238	.409
Embracing 3	.594	-.177	.384
Embracing 4	.708	-.269	.573
Embracing 5	.488	.043	.240
			Total
Sum of Squares (eigenvalues)	3.603	1.184	4.787
Explained variance	36.029	11.838	47.866

Extraction Method: Maximum Likelihood.

Unrotated factor matrix showed that it was not possible to get the optimum factor structures without applying any rotation. However, before that, the items with low communalities were deleted. Although elimination of these items improved KMO value, communalities and explained total variance, rotation was still needed.

Table 35

Promax-Rotated Pattern and Structure Matrices for CEI-II in Study 2: Full set of variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Stretching 1	-.233	.516	.114	.359	.158
Stretching 2	.176	.406	.449	.524	.292
Stretching 3	.112	.620	.529	.695	.489
Stretching 4	.750	-.062	.708	.442	.503
Stretching 5	.457	.216	.602	.523	.388
Embracing 1	.245	-.024	.229	.141	.053
Embracing 2	.196	.491	.526	.623	.409
Embracing 3	.656	-.056	.618	.386	.384
Embracing 4	.846	-.143	.750	.426	.573
Embracing 5	.330	.202	.466	.424	.240
					Total
Sum of Squares (eigenvalues)	3.015	.475			3.490
Explained variance	30.146	4.746			34.892

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

After rotation, factor loadings which are presented in Table 35 were still far from yielding clear sets of structures. Therefore, stretching 4 and 5, and Embracing 1, 2 and 5 were deleted one by one until the structures were clear enough.

Table 36

Promax-Rotated Pattern and Structure Matrices for CEI-II in Study 2: Reduced set of variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Stretching 1	-.170	.456	.037	.379	.167
Stretching 2	.295	.320	.440	.454	.275
Stretching 3	.248	.611	.525	.723	.572
Embracing 3	.939	-.214	.842	.212	.745
Embracing 4	.531	.082	.568	.323	.329
					Total
Sum of Squares (eigenvalues)	1.553	.534			2.087
Explained variance	31.059	10.688			41.746

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

Finally, it was possible to obtain two unidimensional factors. In this respect, Factor 1 (Embracing) had eigenvalue 1.553 and explained 31.059% of the total

variance by itself. Factor 2 (Stretching) had eigenvalue .534 and explained 10.688% of the variance. They both explained 41.746% of the total variance cumulatively. This reduced version of CEI-II showed that it was possible to use the scales as suggested by their authors. In addition, Cronbach's alpha value for Factor 1 was .633 whereas it was .498 for Factor 2. These figures indicated that there were some issues regarding the reliability of Factor 2. On the other hand, high loadings for Factor 1 and 2 indicated the existence of convergent validity.

Table 37

Promax-Rotated Factor Correlation Matrix for CEI-II in Study 2

Factor	1	2
1	1.000	.454
2	.454	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

When Table 37 was examined, it could be seen that, the two factors had moderate correlation, r = .454, which indicated that they were somewhat different from each other. That is, discriminant validity existed.

EFA with Achievement Goal Questionnaire-Revised. Final exploratory factor analysis procedure in Study 2 was applied to test Avoidance (Mastery and Performance) dimension of AGQ-R. The correlation matrix (Appendix 8) revealed that there were correlations of r = .3 or higher. 13 out 15 correlations were significant at .01 (2-tailed) and 1 was significant at .05 (2-tailed) level. Bartlett's Test of Sphericity, approx. chi-square = 599.847, df = 15, $p < .001$, also indicated that there were strong correlations among variables. Kaiser-Meyer-Olkin measure of sampling adequacy value was .717, which indicated that data was large enough to carry out such an analysis. These findings showed that data was suitable for factor analysis.

As in previous analyses, maximum likelihood extraction method was applied and factor extraction was stopped when two factors were obtained. Table 38 gives the unrotated factor loadings, communalities, eigenvalues and percentages of explained total variance by the two factors.

Table 38

Unrotated Factor Matrix for AGQ-R in Study 2

	Factor		Communality
	1	2	
Mastery Avoidance 1	.798	-.463	.851
Mastery Avoidance 2	.668	-.233	.501
Mastery Avoidance 3	.419	-.161	.202
Performance Avoidance 1	.690	.316	.576
Performance Avoidance 2	.595	.600	.714
Performance Avoidance 3	.548	.625	.691
			Total
Sum of Squares (eigenvalues)	2.833	1.412	4.245
Explained variance	47.218	23.534	70.752

Extraction Method: Maximum Likelihood.

* Loadings higher than .3 were given in bold

The matrix provided in Table 38 shows that the lowest communality value belonged to Mastery Avoidance 3 (.202). Although it seemed low, it was not deleted from the analysis not to lose the entire subscale. Unrotated loadings indicated the existence of some factor structures but there were still problems. Mastery Avoidance 1, and Performance Avoidance 1, 2 and 3 were cross loaded on both factors. Thus, Promax rotation was applied. As a result of Promax with Kaiser Normalization rotation, two distinct factors were obtained.

Table 39

Promax-Rotated Pattern and Structure Matrices for AGQ-R in Study 2

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Mastery Avoidance 1	-.093	.953	.266	.918	.851
Mastery Avoidance 2	.074	.677	.328	.705	.501
Mastery Avoidance 3	.031	.436	.195	.448	.202
Performance Avoidance 1	.622	.260	.719	.494	.576
Performance Avoidance 2	.857	-.033	.845	.289	.714
Performance Avoidance 3	.860	-.088	.827	.236	.691
					Total
Sum of Squares (eigenvalues)	2.389	1.145			3.534
Explained variance	39.810	19.090			58.900

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Factor loadings more than .3 have been given in bold

As can be seen from Table 39, rotated factor loadings yielded two distinct structures with high factor loadings. Consequently, it could be argued that the subscales Mastery Avoidance and Performance Avoidance could be used as suggested by their authors.

Above findings also indicate that the unidimensional factor structures existed. Regarding the reliability of these two structures, Cronbach's alpha value for Factor

1 was .833 and for Factor 2 it was .722, both of which were above the suggested levels. In addition, significantly high factor loadings also indicated that convergent validity was achieved.

Table 40

Promax-Rotated Factor Correlation Matrix for AGQ-R in Study 2

Factor	1	2
1	1.000	.376
2	.376	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

Table 40 shows the correlations between factors Mastery Avoidance and Performance Avoidance. Accordingly, relatively low correlation coefficient r = .376 indicated the existence of discriminant validity. In other words, these two factors were different from each other.

Study 3

Data Screening

Type of missing data. As in Study 1 and 2, the same procedures were applied in Study 3 which was the main application. First, the data set was controlled and it was found out that the missing data were not resulted from data entry errors but from participants' not responding the items. That is, all the missing data were unknown in Study 3.

Determination of the extent of missing data. To determine the extent of missing data, the distribution of missing data was examined per case and per variable. According to the findings from Study 1 and Study 2, the number of variables had been reduced in Study 3. After the omission of variables aiming at discovering demographic data, 86 variables were obtained. Out of these 86 variables, 76 variables had at least one missing value. Table 41 shows the distribution of missing values for each variable.

Table 41

Summary of Missing Data per Variable in Study 3

	N	Mean	Std. Deviation	Missing	
				Count	Percent
Autonomy Satisfaction 4	1609	3.49	1.082	12	.7

Autonomy Frustration 3	1610	2.60	1.174	11	.7
Mastery Avoidance 1	1612	3.98	.900	9	.6
DMC5	1613	2.36	1.725	8	.5
Controlled Regulation 2	1613	5.36	1.650	8	.5
Performance Avoidance 1	1614	3.58	1.154	7	.4
Mastery Avoidance 2	1614	4.00	.900	7	.4
DMC8	1615	2.03	1.622	6	.4
Relatedness Satisfaction 1	1615	3.63	1.005	6	.4
Autonomy Frustration 2	1615	2.82	1.205	6	.4
Autonomy Frustration 4	1615	3.06	1.233	6	.4
Relatedness Satisfaction 4	1615	4.10	.843	6	.4
L2B Entity	1615	2.72	1.522	6	.4
Autonomous Regulation 1	1615	5.12	1.686	6	.4
Controlled Regulation 5	1615	3.96	1.893	6	.4
Autonomous Regulation 3	1615	5.36	1.532	6	.4
DMC4	1616	2.42	1.849	5	.3
Mastery Approach 1	1616	4.18	.799	5	.3
Performance Avoidance 3	1616	3.54	1.211	5	.3
Relatedness Frustration 1	1616	1.91	.999	5	.3
Competence Satisfaction 1	1616	3.91	.950	5	.3
Competence Satisfaction 2	1616	3.72	.957	5	.3
Autonomy Satisfaction 3	1616	3.64	.983	5	.3
Relatedness Frustration 3	1616	2.03	1.079	5	.3
Relatedness Frustration 4	1616	2.55	1.115	5	.3
Controlled Regulation 1	1616	2.21	1.655	5	.3
DMC3	1617	2.29	1.753	4	.2
DMC10	1617	2.68	1.917	4	.2
Performance Approach 2	1617	3.44	1.169	4	.2
Performance Avoidance 2	1617	3.63	1.107	4	.2
Mastery Avoidance 3	1617	4.08	.839	4	.2
Relatedness Satisfaction 2	1617	3.80	1.025	4	.2
Embracing 2	1617	3.87	.944	4	.2
Stretching 3	1617	3.84	.956	4	.2
Psychological Ownership 3	1617	2.93	1.188	4	.2
L2B Incremental	1617	4.87	1.162	4	.2
Autonomous Regulation 2	1617	5.20	1.593	4	.2
Mastery Approach 3	1618	4.13	.850	3	.2
Performance Approach 3	1618	3.31	1.191	3	.2
Competence Frustration 1	1618	2.46	1.152	3	.2
Autonomy Satisfaction 2	1618	3.58	.966	3	.2
Relatedness Frustration 2	1618	2.11	1.070	3	.2
Relatedness Satisfaction 3	1618	3.96	.858	3	.2
Competence Satisfaction 3	1618	3.80	.981	3	.2
Embracing 1	1618	2.33	1.320	3	.2
L2B Incremental	1618	4.45	1.263	3	.2
Autonomous Regulation 5	1618	4.17	1.653	3	.2
DMC2	1619	2.62	1.857	2	.1
DMC7	1619	2.17	1.697	2	.1
Performance Approach 1	1619	3.30	1.104	2	.1
Competence Frustration 2	1619	2.59	1.131	2	.1
Competence Frustration 3	1619	2.51	1.228	2	.1
Embracing 3	1619	3.32	1.170	2	.1
Embracing 5	1619	3.22	1.253	2	.1
Psychological Ownership 5	1619	2.77	1.150	2	.1
Scarcity 3	1619	4.03	1.136	2	.1
L2B Entity	1619	2.86	1.425	2	.1
L2B Incremental	1619	4.99	1.133	2	.1
Controlled Regulation 3	1619	4.08	1.911	2	.1

Controlled Regulation 4	1619	3.71	1.902	2	.1
Autonomous Regulation 4	1619	5.04	1.716	2	.1
Controlled Regulation 6	1619	3.96	1.980	2	.1
DMC1	1620	2.60	1.828	1	.1
DMC9	1620	2.32	1.760	1	.1
DMC11	1620	2.76	1.961	1	.1
DMC12	1620	2.33	1.757	1	.1
Mastery Approach 2	1620	4.36	.774	1	.1
Autonomy Satisfaction 1	1620	3.62	1.066	1	.1
Competence Satisfaction 4	1620	3.84	.894	1	.1
Stretching 1	1620	3.73	.891	1	.1
Stretching 5	1620	3.46	1.050	1	.1
Psychological Ownership 1	1620	3.03	1.114	1	.1
Psychological Ownership 2	1620	2.95	1.078	1	.1
Psychological Ownership 4	1620	2.68	1.008	1	.1
Controlled Regulation 7	1620	2.80	1.934	1	.1

When Table 41 was examined, it could be seen that Autonomy Satisfaction 3 (12 missing values) and Autonomy Frustration 3 (11 missing values) had the highest number of missing values. However, when compared to the number of participants, none of the variables had missing values more than 8%, which was below the suggested limits (Hair et al., 2013). In the next step, missing values were investigated in case-base.

Table 42

Amount of Missing Data per Case in Study 3

Number of Missing Data per Case	Number of Cases	Percent
0	1416	87.35
1	153	9.44
2	36	2.22
3	9	0.56
4	6	0.37
7	1	0.06
Total	1621	100

As can be seen from Table 42, the number of missing data per case was not more than 15% of the cases. 87.35% of all cases had no missing data and 9.44% had only one missing value. In sum, 12.65% of all cases had at least one missing value. As none of the single variables and cases had significant amount of missing data, there was no problem in proceeding for further analyses. The findings from the above examinations of the data revealed that the extent of the missing data was low enough to go on with the fourth step in the process.

Randomness of the missing data. Regarding the randomness of the data, Little's MCAR test was carried out. It was found out that the data was not missing completely at random (MAR), Little's MCAR test: Chi-Square = 10326.446, df = 9194, Sig = .000. This finding meant that it was possible to use the estimated EM means obtained from the test. However, to enable comparison, regression method as multiple imputation method (Schaffer, 1999; McKnight et al., 2007) was carried out as well. Another suggestion for the replacement values for the missing data is put forward by Gaskin (2018). Gaskin (2018) argues that with Likert type items, replacing missing values with mean scores is not appropriate. Instead, to impute the missing values, it would be more appropriate to use the surrounding values of other indicators for the latent factors. When the data set was inspected in terms of the missing values, it was seen that there were some cases in which the participants indicated that they had not experienced any long-term motivation. In such cases the values preferred would be zero (0), which means that when the missing values in these cases were replaced by the mean scores obtained from either as a result, the EM imputation or Multiple Regression imputation could be nonsense. So, in the scope of this study, replacement of the missing values with the mode value for that series in the case was preferred. Table 43 provides the mean scores and standard deviations as a result of above-mentioned imputation methods for the variables which had been used in EFA procedures in Study 3. The findings suggest that almost no changes could be observed due to different methods. Thus, it would be correct to assume that replacing the missing values with series modes was acceptable.

Table 43

Means and Standard Deviations after the application of Different Missing Data Replacement Procedures

Variable	All Data		Data Replacement Method					
			EM Imputation		Regression Imputation		Series Mode	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Autonomy Satisfaction 4	3.49	1.082	3.49	1.082	3.50	1.083	3.50	1.081
Mastery Avoidance 1	3.98	.899	3.98	.900	3.98	.900	3.98	.899
DMC5	2.36	1.725	2.36	1.722	2.36	1.723	2.36	1.722
Performance Avoidance 1	3.58	1.154	3.58	1.153	3.58	1.153	3.58	1.152
Mastery Avoidance 2	4.00	.900	4.01	.899	4.01	.899	4.00	.899
Relatedness Satisfaction 4	4.10	.843	4.10	.844	4.10	.842	4.09	.846
Relatedness Satisfaction 1	3.63	1.005	3.63	1.005	3.63	1.004	3.63	1.005
DMC8	2.03	1.622	2.03	1.621	2.03	1.622	2.03	1.621
Competence Satisfaction 2	3.72	.957	3.72	.957	3.72	.957	3.72	.954

Competence Satisfaction 1	3.91	.950	3.91	.950	3.91	.950	3.91	.949
Autonomy Satisfaction 3	3.64	.983	3.64	.983	3.64	.982	3.64	.983
Performance Avoidance 3	3.54	1.211	3.54	1.211	3.54	1.210	3.54	1.210
DMC4	2.42	1.849	2.43	1.848	2.43	1.848	2.43	1.848
Embracing 2	3.87	.944	3.87	.944	3.87	.945	3.87	.944
Relatedness Satisfaction 2	3.80	1.025	3.80	1.025	3.79	1.027	3.80	1.025
Performance Avoidance 2	3.63	1.107	3.63	1.108	3.63	1.107	3.63	1.109
Mastery Avoidance 3	4.08	.839	4.08	.839	4.08	.839	4.08	.839
DMC10	2.68	1.917	2.68	1.918	2.68	1.918	2.68	1.918
DMC3	2.29	1.753	2.30	1.753	2.30	1.753	2.29	1.753
Embracing 1	2.33	1.320	2.33	1.320	2.34	1.320	2.33	1.320
Competence Satisfaction 3	3.80	.981	3.80	.980	3.80	.980	3.80	.980
Relatedness Satisfaction 3	3.96	.858	3.96	.858	3.96	.857	3.96	.858
Autonomy Satisfaction 2	3.58	.966	3.58	.966	3.58	.965	3.58	.965
Scarcity 3	4.03	1.136	4.03	1.136	4.03	1.136	4.03	1.136
Embracing 5	3.22	1.253	3.22	1.253	3.22	1.253	3.22	1.252
Embracing 3	3.32	1.170	3.32	1.170	3.32	1.170	3.32	1.170
DMC7	2.17	1.697	2.17	1.696	2.17	1.696	2.17	1.696
DMC2	2.62	1.857	2.63	1.856	2.62	1.856	2.62	1.856
Stretching 5	3.46	1.050	3.46	1.050	3.46	1.050	3.46	1.050
Competence Satisfaction 4	3.84	.894	3.84	.894	3.84	.894	3.84	.984
Autonomy Satisfaction 1	3.62	1.066	3.62	1.066	3.62	1.066	3.62	1.066
DMC12	2.33	1.757	2.33	1.757	2.33	1.757	2.33	1.757
DMC11	2.76	1.961	2.76	1.960	2.76	1.960	2.76	1.960
DMC9	2.32	1.760	2.32	1.760	2.32	1.760	2.32	1.760
DMC1	2.60	1.828	2.60	1.828	2.60	1.828	2.60	1.828

Outliers. Tabachnick and Fidell (2013) argue that procedural errors are among the reasons of outliers. To eliminate this kind of outliers, the data was investigated by conducting frequency analysis and extreme values in each variable was sought out. When such values had been found, participants' papers, which had been given a unique ID number during data entry process, were checked and errors were fixed. Another type of outliers might result from unengaged responses by participants, which means the cases in which participants answered almost every item somewhat agree or disagree. As Gaskin (2018) suggested, first standard deviation values for each case were calculated because if a case has a standard deviation of 0, it means that all the responses are the same. Findings indicated that many of cases had standard deviation value above 0. However, two cases (cases 357 and 371) had very low SDs. Thus, for these cases, visual inspection was carried out and it was seen that the participants had chosen the same option in every case. As a result, cases numbered 357 and 371 were removed due to their being unengaged. Finally, there were 1619 complete cases ready for the analyses.

Although data should be investigated in terms of the existence of univariate, bivariate and multivariate outliers (Hair et al., 2013), as Zijlstra (2009) argues, the methods developed for this purpose aim at investigating continuous data and cannot

be employed for Likert type questionnaire data (p.89). Moreover, Hair et al. (2013) also state that to enable the generalizability of the findings, outliers could be retained (p. 65). Therefore, it was decided not to carry out any outlier detection method.

Statistical Assumptions

Normality. Normally the next step following the missing value diagnosis and replacement procedures would be spotting the univariate and multivariate outliers that had the potential to have effect on the outcome due to their extreme values and might have resulted in biased conclusions (Tabachnick & Fidell, 2013). However, current study only employed ordinal Likert type data, which were not suitable for these kinds of analyses. Riani, Torti and Zani (2011) point out that although there might be some special cases in which the frequency distribution of a variable could reveal univariate outliers, it is difficult to detect outliers in ordinal variables (p.159). Because ordinal variables link to ranks that range from predetermined values (from 1 to n), it may not be possible to define them as outliers. Similarly, Gaskin (2019) pointed out that there are not any outliers in Likert-Scales. Because when a participant answer at the extreme, which can only be either 1 or 5, it may not be appropriate to identify that answer as an outlier.

Most multivariate analyses are based on several assumptions. Hair et al, (2013) point out that four of these assumptions have potential effects on every univariate and multivariate statistical techniques: normality, homoscedasticity, linearity and absence of correlated errors (pp. 69-74). Perhaps the most influential of these four assumptions is normality. That is, the data need to be normally distributed across variables. When the individual variables are tested, it is called as univariate normality. On the other hand, when the combination of two or more variables is tested, it is called multivariate normality. Hair et al. (2013) suggest that testing the univariate normality of all variables would be satisfactory in many cases (p. 69). Since normal distribution of the data is closely related to the shape of distribution, this can be tested by examining the measures of Skewness and Kurtosis. Hair et al. (2013) state that Skewness values outside the range of -1 to +1 cause problems in terms of normal distribution of the data. However, Tabachnick and Fidell (2013) argues the use of ±1.5 as the tolerable range. George and Mallery (2013) obtain looser limit and state that a value between ±2 is also acceptable.

Similarly, Sposito, Hand, and Skarpness (1983) argue that it should be between ±2.2. Based on these published thresholds, it could be said that the data achieved univariate normality. Table 44 shows the Skewness and Kurtosis values for the variables used in scope of the current study.

Table 44

Skewness and Kurtosis Values for Study 3

	N			Std. Error of		Std. Error of
Item	Valid	Missing	Skewness	Skewness	Kurtosis	Kurtosis
Embracing 1	1619	0	.614	.061	-.823	.122
DMC8	1619	0	.110	.061	-1.254	.122
Psychological Ownership 3	1619	0	.095	.061	-.985	.122
Psychological Ownership 4	1619	0	.080	.061	-.324	.122
Psychological Ownership 5	1619	0	.068	.061	-.710	.122
DMC7	1619	0	.022	.061	-1.318	.122
DMC6	1619	0	-.049	.061	-1.263	.122
DMC3	1619	0	-.090	.061	-1.391	.122
Psychological Ownership_3	1619	0	-.095	.061	-.985	.122
DMC4	1619	0	-.113	.061	-1.436	.122
Stretching 2	1619	0	-.115	.061	-.262	.122
DMC9	1619	0	-.120	.061	-1.403	.122
Embracing 5	1619	0	-.125	.061	-.971	.122
Psychological Ownership 2	1619	0	-.128	.061	-.605	.122
DMC12	1619	0	-.141	.061	-1.384	.122
Stretching 4	1619	0	-.145	.061	-.682	.122
Psychological Ownership 1	1619	0	-.204	.061	-.670	.122
Embracing 4	1619	0	-.206	.061	-.651	.122
Embracing 3	1619	0	-.228	.061	-.811	.122
DMC5	1619	0	-.241	.061	-1.326	.122
Stretching 1	1619	0	-.259	.061	-.198	.122
Stretching 5	1619	0	-.276	.061	-.533	.122
DMC10	1619	0	-.386	.061	-1.442	.122
DMC2	1619	0	-.407	.061	-1.384	.122
DMC11	1619	0	-.425	.061	-1.440	.122
DMC1	1619	0	-.440	.061	-1.367	.122
Autonomy Satisfaction 2	1619	0	-.446	.061	.071	.122
Autonomy Satisfaction 4	1619	0	-.464	.061	-.327	.122
Competence Satisfaction 2	1619	0	-.560	.061	.123	.122
Stretching 3	1619	0	-.604	.061	-.021	.122
Competence Satisfaction 4	1619	0	-.606	.061	.410	.122
Embracing 2	1619	0	-.623	.061	-.046	.122
Relatedness Satisfaction 1	1619	0	-.640	.061	.151	.122
Autonomy Satisfaction 3	1619	0	-.656	.061	.247	.122
Competence Satisfaction 3	1619	0	-.665	.061	.221	.122
Autonomy Satisfaction 1	1619	0	-.673	.061	-.036	.122
Performance Avoidance 3	1619	0	-.684	.061	-.477	.122
Performance Avoidance 1	1619	0	-.706	.061	-.341	.122
Scarcity 2	1619	0	-.716	.061	-.968	.122
Performance Avoidance 2	1619	0	-.785	.061	-.082	.122
Competence Satisfaction 1	1619	0	-.804	.061	.518	.122
Scarcity 5	1619	0	-.816	.061	-.588	.122
Relatedness Satisfaction 3	1619	0	-.864	.061	.988	.122
Relatedness Satisfaction 2	1619	0	-.917	.061	.529	.122
Relatedness Satisfaction 4	1619	0	-1.044	.061	1.437	.122

Mastery Avoidance 3	1619	0	-1.048	.061	1.570	.122
Mastery Avoidance 1	1619	0	-1.132	.061	1.605	.122
Scarcity 1	1619	0	-1.177	.061	.791	.122
Scarcity 3	1619	0	-1.198	.061	.761	.122
Mastery Avoidance 2	1619	0	-1.241	.061	1.966	.122
Scarcity 4	1619	0	-1.331	.061	.903	.122

As can be seen in Table 44, many of the variables had Skewness values less than ±1 and 7 variables had Skewness values between ±2. Regarding the Kurtosis, all variables had Kurtosis values between ±2. According to these findings, it could be stated that all of the Skewness and Kurtosis values were in range of suggested limits and thus it could be assumed that the data was normally distributed.

Another factor affecting the normality of the data is the sample size. Hair et al. (2013) state that sample size had significant effect on normality. As the sample size increases, the detrimental effect of sample size decreases because large sample sizes tend to increase statistical power, which in turn reduces the sampling error (p.75). In this respect, when the sample size of this study, which is 1619, was taken into consideration, it could be said that the sample size was large enough to have a normal distribution.

Linearity. To be able to carry out factor analysis, the data need to have a linear distribution. If not, curvilinearity occurs. That is, there should not be a lot of differences in Skewness values of variables. To check linear distribution, Scatterplots can be controlled (Tabachnick & Fidell, 2013). However, for 50 variables, it was almost not possible to check all the scatterplots. In this case, sample scatterplot which includes the highest negative and highest positive Skewness values can be examined (Tabachnick & Fidell, 2013; p.657). Figure 12 shows the sample scatterplots.

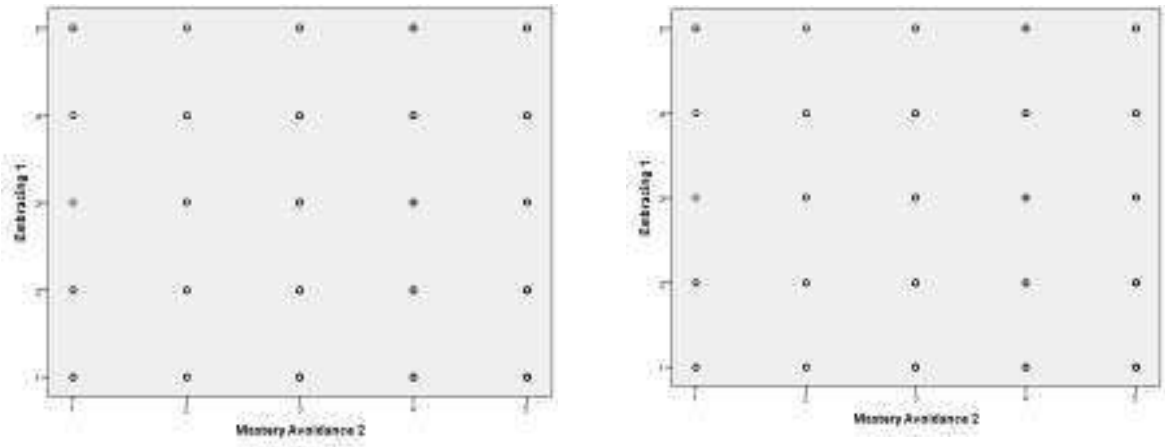


Figure 12. Scatterplots for linearity check among variables.

The highest positive Skewness values were .614 for Embracing 1 and .110 for DMC 8. On the other hand, the highest negative Skewness values were -1.241 for Mastery Avoidance 2 and -1.331 for Scarcity. As can be seen in Figure 12, when the scatterplots were examined, it was not possible to talk about the existence of linear relationships between these variables. However, there was also no evidence showing the existence of curvilinearity. Therefore, as in the previous studies no actions were taken in terms of data transformation.

Exploratory Factor Analyses

In Study 1 and Study 2 Octalysis survey which included the adapted versions of The Curiosity and Exploration Inventory 2 (CEI-II) by Kashdan et al. (2009), Basic Psychological Need Satisfaction and Frustration Scale-General Measure (BPNSF-General) by Chen et al. (2015), Achievement Goal Questionnaire Revised (AGQ-R) by Elliot and Murayama (2008), Directed Motivational Currents Disposition Questionnaire (DMCDQ) by Muir (2016), Psychological Ownership Questionnaire (POQ) by Van Dyne and Pierce (2004), and Susceptibility to Persuasion Scale (STPS) by Cialdini (2012) were utilized. Study 1, which was carried out with 90 participants, and Study 2, which was carried out with 278 participants, revealed that Octalysis survey could be used as suggested by their authors providing the problems occurred were solved.

These two piloting studies showed that sample size plays a great role in this sense. Thus, in the scope of Study 3, new adaptations offered by the participants and experts had been done and the new version of survey was applied with 1619 participants. As was in the previous studies, the same procedures were applied and different EFAs were conducted to test the scales and questionnaires. Before proceeding to the statistical analyses, the data set was trimmed and only the dimensions which cover the eight core drives of Octalysis framework were kept. That is, 50 variables covering 10 dimensions were obtained.

Unit of Analysis. Factor analysis is not used just to determine the structures among variables. There are two types of factor analysis. The one which is applied to correlation matrix of variables and aim at identifying the latent dimensions underlying a set of variables is called R factor analysis. The other which is applied to correlation matrix of individual respondents and aim at linking or reducing large

number of participants into different groups is called Q factor analysis (Hair et al., 2013; p.96). Since the aim of this study is to investigate the latent dimensions of variables, R factor analysis is preferred.

Assumptions in Factor Analysis. After deleting the cases with unengaged responses, there were 1619 cases in Study 3. In addition, cases in which participants stated they had never experienced DMC were also eliminated. As a result, a data set with 1150 cases and 50 variables was obtained. Hair et al. (2013) argue that to conduct factor analysis, sample size should have 5 to 1 ratio and preferably 10 to 1 ratio (p.100). That is, there should be at least ten cases per variable. In addition, they state that sample size should not be lower than 50 and should be larger than 100. Accordingly, sample size (n = 1150) in Study 3 was above the suggested limits and had a ratio over 20 to 1. Thus, it was decided to proceed with the EFA procedures.

As in Study 1 and Study 2, first DMC, Scarcity and Psychological Ownership dimensions were factor analyzed. Then, items in BPNFS-General, CEI-II and AGQ-R were subjected to the exploratory factor analyses, respectively. In each EFA procedure, Maximum Likelihood Extraction method (MLE) was applied. The calculation method used in MLE maximizes the probability of the observed correlations (Tabachnick & Fidell, 2013; p.641). Since MLE method yields indexes of the goodness of fit, statistical significance levels for factor loadings, and correlations among factors (Cudeck & O'Dell, 1994) and since MLE is the extraction method used in AMOS software which is used for SEM procedures, it was preferred for this study. In this way, it was aimed to get more consistent results for further SEM analyses.

EFA with DMC, Scarcity and Psychological Ownership. Before conducting factor analysis, first the suitability of the data for factoring was controlled. In this respect, as Pallant (2011) suggested, a correlation matrix (Appendix 9) was produced. When the correlation matrix was investigated, it was found out that there were many correlations of .3 or above. In addition, out of 250 correlations, 185 were significant at .01 (2-tailed) and 11 were significant at .05 (2-tailed) level. These findings indicate that further factor analysis can be carried out. Another way of determining whether variables have significant correlations is the Bartlett's Test of Sphericity (Hair et al., 2013). Bartlett's test, approx. chi-square = 6640.746, df = 231,

p < .001, showed that correlations among variables in Study 3 were significant at .001 level. Next, the adequacy of the sampling was checked. According to Hair et al. (2013) measure of sampling adequacy (MSA) should be above .80 for best (p.102). Kaiser-Meyer-Olkin Measure of sampling adequacy value was .843 for the current factor analysis. Based on this finding, it was concluded that data were suitable for factor analysis.

Factor analysis with MLE was conducted with a priori criterion (Hair et al., 2013) and factor extraction was stopped when three factors were obtained.

Table 45

Unrotated Factor Matrix for DMC, Psychological Ownership and Scarcity in Study 3

	Factor			Communality
	1	2	3	
DMC1	.315	.505	-.248	.416
DMC2	.234	.473	-.097	.288
DMC3	.161	.386	.015	.175
DMC4	-.143	-.015	.352	.144
DMC5	.308	.447	-.202	.336
DMC6	.320	.553	-.303	.500
DMC7	.216	.483	.143	.301
DMC8	.238	.363	.100	.198
DMC9	.137	.395	.032	.176
DMC10	.211	.526	.451	.524
DMC11	.174	.509	.340	.404
DMC12	.342	.398	-.366	.410
Psychological Ownership 1	.898	-.195	.029	.845
Psychological Ownership 2	.790	-.171	.027	.655
Psychological Ownership 3	.695	-.171	-.052	.515
Psychological Ownership 4	.170	-.050	-.054	.034
Psychological Ownership 5	.819	-.162	.032	.698
Scarcity 1	.231	.215	.041	.101
Scarcity 2	.247	.138	.113	.093
Scarcity 3	.160	.218	.102	.083
Scarcity 4	.125	.164	.105	.054
Scarcity 5	.296	.176	.118	.133
				Total
Sum of Squares (eigenvalues)	4.652	2.603	1.618	8.873
Explained variance	21.144	11.832	7.357	40.333

Extraction Method: Maximum Likelihood.

When unrotated initial factor loadings which are presented in Table 45 were investigated, without rotation it was hard to talk about the existence of clear set of structures. Items aimed at measuring DMC, DMC 1, DMC 5, DMC 6, DMC 10, DMC 11 and DMC 12 cross loaded onto different factors and did not form a clear factor structure. Items for Psychological Ownership seemed to load onto a single factor

and apart from Psychological Ownership 4 they all had significantly high loadings. Thus, it could be stated that unrotated factor loadings for Psychological Ownership indicated a factor. However, the same cannot be said for the items aiming at the measurement of Scarcity. For this factor, there were no items loaded significantly high. When the communalities were investigated, it was seen that there were many items which had too low loadings. Thus, it was decided to apply Promax rotation method. But before that, Psychological Ownership 4 was deleted because of its low communality value. No action was taken regarding Scarcity at this stage.

In the next step, Promax Kaiser Normalization rotation was applied. It was seen that after the elimination of Psychological Ownership 4, Factor 1 (Psychological Ownership) had significantly high loadings. However, there were problems regarding DMC and Scarcity items. In terms of DMC, the items loaded onto two factors. Therefore, DMC 4 which cross loaded onto two factors and DMC 3 which had loading less than .3 were deleted. Then until reaching a single factor for DMC, items 8, 9 and 11 were deleted separately. At the end, a Scarcity structure with seven items was obtained. Obtaining clear structures for DMC and Psychological Ownership resulted in significant improvements in Scarcity. At the end of the item deletion process, all the scarcity items significantly loaded onto a single factor which is Factor 3. Similar improvements were observed regarding the communalities. Yet, there were items with low communalities. Thus, DMC 10 and Scarcity 2 were deleted. This caused the factor loadings and communalities to increase. At the end, although DMC 7 and Scarcity 4 had communalities lower than the suggested level (Child, 2006), they were retained for theoretical considerations. Tables 46 and 47 provide information for full and reduced sets of variables after Promax rotation for Factor 1 (Psychological Ownership), Factor 2 (DMC) and Factor 3 (Scarcity) in terms of their factor loadings, communalities, sum of square values, and percentages of explained variance.

Table 46

Promax Rotated Pattern and Structure Matrices for DMC, Psychological Ownership and Scarcity in Study 3: Full set of variables

	Pattern Matrix*			Structure Matrix*			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
DMC1	-.001	.135	.566	.182	.421	.634	.416
DMC2	-.040	.247	.379	.111	.431	.492	.288
DMC3	-.048	.286	.205	.062	.381	.337	.175

DMC4	-.078	.305	-.400	-.134	.088	-.268	.144
DMC5	.028	.141	.487	.190	.392	.566	.336
DMC6	-.028	.114	.650	.175	.438	.700	.500
DMC7	-.032	.479	.127	.092	.538	.361	.301
DMC8	.044	.365	.113	.143	.430	.310	.198
DMC9	-.072	.304	.189	.037	.386	.322	.176
DMC10	-.020	.800	-.179	.077	.706	.220	.524
DMC11	-.059	.677	-.074	.045	.629	.252	.404
DMC12	.063	-.041	.640	.234	.294	.636	.410
Psychological Ownership 1	.914	.045	-.013	.919	.207	.266	.845
Psychological Ownership 2	.804	.042	-.013	.808	.184	.234	.655
Psychological Ownership 3	.708	-.049	.059	.716	.111	.232	.515
Psychological Ownership 4	.172	-.056	.054	.177	.003	.073	.034
Psychological Ownership 5	.826	.057	-.009	.834	.205	.251	.698
Scarcity 1	.105	.213	.098	.172	.282	.235	.101
Scarcity 2	.167	.236	-.016	.206	.258	.150	.093
Scarcity 3	.047	.262	.026	.103	.283	.172	.083
Scarcity 4	.043	.225	-.010	.082	.228	.116	.054
Scarcity 5	.193	.273	.005	.245	.311	.197	.133
							Total
Sum of Squares (eigenvalues)	3.569	2.645	.870				7.084
Explained variance	16.225	12.022	3.954				32.200

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Loadings higher than .3 were given in bold

Table 47

Promax-Rotated Pattern and Structure Matrices for DMC, Psychological Ownership and Scarcity in Study 3: Reduced set of variables

	Pattern Matrix*			Structure Matrix*			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
DMC1	-.002	.669	-.020	.200	.660	.240	.436
DMC2	-.048	.537	.039	.130	.538	.235	.292
DMC5	.024	.575	.012	.206	.587	.243	.346
DMC6	-.028	.697	.015	.194	.694	.279	.482
DMC7	-.015	.408	.015	.116	.409	.170	.168
DMC12	.074	.597	-.052	.245	.599	.202	.365
Psychological Ownership 1	.925	-.021	.016	.923	.274	.279	.853
Psychological Ownership 2	.803	-.015	.023	.805	.244	.252	.649
Psychological Ownership 3	.721	.017	-.039	.715	.227	.178	.512
Psychological Ownership 5	.832	.016	-.012	.833	.271	.237	.695
Scarcity 1	.012	.057	.496	.175	.255	.522	.276
Scarcity 3	-.063	-.010	.576	.102	.195	.554	.311
Scarcity 4	-.036	-.021	.427	.083	.135	.409	.169
Scarcity 5	.094	-.005	.530	.247	.231	.556	.317
							Total
Sum of Squares (eigenvalues)	3.219	1.874	.778				5.871
Explained variance	22.990	13.385	5.560				41.935

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Loadings higher than .3 were given in bold

According to Table 47, it can be stated that a clear three-factor structure was obtained at the end of fifth respecification. In addition, since all the items loaded only one factor, it can be argued that these factors were unidimentional. Next, as Hair et

al. (2013) and Gaskin (2018) suggested, the reliability of these structures were examined. Cronbach's Alpha scores for Factors 1, 2 and 3 were .888, .746 and .580, respectively. Accordingly, it could be argued that Factor 1 (Psychological Ownership) and Factor 2 (DMC) had reliable structures whereas Factor 3 (Scarcity) had some problems regarding its reliability.

The validity of these structures, it could be said that they all had convergent validity since all the items loaded onto a single factor significantly. Regarding their discriminant validity, factor correlation matrix (Table 48) for Factors 1, 2 and 3 was examined.

Table 48

Promax-Rotated Factor Correlation Matrix for DMC, Psychological Ownership and Scarcity in Study 3

Factor	1	2	3
1	1.000	.311	.292
2	.311	1.000	.390
3	.292	.390	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

Hair et al. (2013) and Farrell and Rudd (2009) point out that to talk about discriminant validity, factors need to have low correlations. When Table 48 is examined, it can be seen that Factor 1, Factor 2 and Factor 3 have low correlations. This finding indicated that these structures had discriminant validity. That is, they measured different latent structures.

EFA with Basic Psychological Needs Satisfaction and Frustration General Scale. For the second EFA in Study 3, items aimed at measuring Competence Satisfaction, Relatedness Satisfaction and Autonomy Satisfaction subscales of BPNSF-General were used. First, the suitability of data set for such kind of analysis was checked. Bivariate correlation matrix (Appendix 10) indicated that there were many correlations of .3 or higher. In addition, all the correlations were significant at .01 level (2-tailed). Bartlett Test of Sphericity, approx. chi-square = 3331.525, df = 66, $p < .001$, also indicated that correlations among these items were significant. In terms of adequacy of sample size, Kaise-Meyer-Olkin measure of sampling adequacy value, .857 indicated that the data was large enough to conduct such analysis.

First, unrotated factor extraction was applied by using Maximum likelihood extraction method. When initial eigen values and unrotated factor structures were investigated, it was seen that Factor 1 (eigenvalue = 3.956), Factor 2 (eigenvalue = 1.571) and Factor 3 (1.096) explained 55.187% of the total variance. But it was not possible to obtain clear factor structures without applying any rotation. Table 49 provides information regarding unrotated factor structures for Factors 1, 2 and 3.

Table 49

Unrotated Factor matrix for BPNSF-General in Study 3

	Factor			Communality
	1	2	3	
Autonomy Satisfaction 1	.485	.071	-.202	.281
Autonomy Satisfaction 2	.600	-.140	-.273	.454
Autonomy Satisfaction 3	.664	-.010	-.376	.582
Autonomy Satisfaction 4	.542	-.041	-.241	.353
Relatedness Satisfaction 1	.366	.332	.109	.256
Relatedness Satisfaction 2	.297	.385	.067	.241
Relatedness Satisfaction 3	.450	.601	.100	.573
Relatedness Satisfaction 4	.329	.398	.091	.275
Competence Satisfaction 1	.587	-.198	.265	.454
Competence Satisfaction 2	.623	-.152	.067	.416
Competence Satisfaction 3	.639	-.222	.166	.486
Competence Satisfaction 4	.628	-.229	.320	.550

Extraction Method: Maximum Likelihood.

As can be seen from Table 49 communality values for all items were above the suggested limits (Chil, 2006). But all items loaded onto only one factor, which was not expected. Therefore, Promax rotation was applied.

Table 50

Promax-Rotated Pattern and Structure Matrices for BPNSF-General in Study 3: Full set of variables

	Pattern Matrix*			Structure Matrix*			Communality
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	
Autonomy Satisfaction 1	-.019	.480	.120	.333	.519	.321	.281
Autonomy Satisfaction 2	.091	.648	-.096	.476	.665	.216	.454
Autonomy Satisfaction 3	-.066	.795	.020	.455	.761	.340	.582
Autonomy Satisfaction 4	.036	.568	.004	.405	.593	.262	.353
Relatedness Satisfaction 1	.094	-.004	.468	.256	.258	.499	.256
Relatedness Satisfaction 2	-.017	.011	.491	.163	.213	.490	.241
Relatedness Satisfaction 3	-.037	.017	.762	.242	.322	.757	.573
Relatedness Satisfaction 4	.017	-.006	.520	.196	.230	.524	.275
Competence Satisfaction 1	.690	-.039	.023	.673	.417	.249	.454
Competence Satisfaction 2	.470	.226	.012	.620	.535	.275	.416

Competence Satisfaction 3	.625	.116	-.023	.692	.510	.247	.486
Competence Satisfaction 4	.788	-.086	.019	.739	.431	.259	.550
							Total
Sum of Squares (eigenvalues)	3.398	.968	.554				4.9200
Explained variance	28.317	8.068	4.615				41.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Loadings higher than .3 were given in bold

Table 50 shows factor loadings, sum of squares, explained variance in terms of pattern and structure matrices for Competence Satisfaction (Factor 1), Autonomy Satisfaction (Factor 2) and Relatedness Satisfaction (Factor 3) subscales of BPNSF-General.

Table 51

Promax-Rotated Factor Correlation Matrix for BPNSF-General in Study 3

Factor	1	2	3
1	1.000	.646	.352
2	.646	1.000	.432
3	.352	.432	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

Findings indicated that it was possible to obtain clear and unidimensional sets of structures. Although, Factors 2 and 3 had low eigenvalues, no action was taken since there were theoretically three factors. In addition, these three factors explained 41% of the total variance. Moreover, Cronbach's alpha was .774 for Factor 1, .647 for Factor 2 and .723 for Factor 3. This indicates that Factor 1 and 3 had acceptable reliability values whereas the reliability value for Factor 2 was slightly lower than the suggested level (Hair et al., 2013).

Regarding the discriminant validity of the three factors, factor correlation matrix in Table 51 was examined. It was found out that Factors 1 and 2 had relatively high correlation which indicated that they somehow measure the same latent construct. Thus, it could be said that in terms of discriminant validity, these three factors had problems. Yet, since these factors had high enough factor loadings, no action was taken to respecify the structures.

EFA with the Curiosity and Exploration Inventory 2. The third EFA procedure was applied for CEI-II. First, the correlation matrix was examined and it was seen that there were many correlations over .3 and almost all of these correlations were significant at .01 (2-tailed) level. Only one correlation between Embracing 1 and Stretching 1 was insignificant. Bartlett's Test of Sphericity, approx.

chi-square = 3002,978, df = 45, p< .001, indicated the significance of correlations among variables as well. Kaiser-Meyer-Olkin Measure of Sampling Adequacy test, .896, showed that the sample size was large enough to carry out factor analysis with this data.

Table 52

Unrotated Factor Matrix for CEI-II in Study 3

	Factor		Communality
	1	2	
Stretching 1	.459	-.306	.304
Stretching 2	.632	.031	.400
Stretching 3	.674	-.281	.534
Stretching 4	.666	.104	.455
Stretching 5	.662	-.141	.458
Embracing 1	.296	.287	.170
Embracing 2	.637	-.165	.433
Embracing 3	.608	.310	.465
Embracing 4	.627	.284	.474
Embracing 5	.474	.022	.225
			Total
Sum of Squares (eigenvalues)	4.002	1.137	5.139
Explained variance	40.017	11.368	51.385

Extraction Method: Maximum Likelihood.

Maximum likelihood extraction method yielded the unrotated factor matrix which was displayed in Table 52 Accordingly, all the items loaded onto a single factor and it was not possible to talk about the existence of clear structures. Factor 1 had eigenvalue of 4.002 and explained 40.017% of the total variance. On the other hand, the other factor had eigenvalue of 1.137 and explained 11.368% of the variance. Communality values were high enough except the one for Embracing 1. These findings indicated the necessity of rotation. But, before refining the factor analysis with Promax rotation, Embracing 1 with communality .170 was deleted.

After Promax rotation, it was found out that there were still some problems in terms of the factor structures. Stretching 2 significantly cross loaded onto the two factors whereas Embracing 5 did not load significantly onto neither of the factors. So these items were eliminated and factor analysis was rerun. When pattern matrix was examined, it was seen that Stretching 4 and Embracing 2 loaded onto wrong factor structures. Therefore, in the second respecification, these two variables were deleted. Final model yielded two clear structures with three items in each. Tables 53 and 54 provide pattern and structure matrices for both full and reduced sets of

variables together with communality values and the eigenvalues, percentage of explained variance each factor had.

Table 53

Promax-Rotated Pattern and Structure Matrices for CEI-II in Study 3: Full Sets of Variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Stretching 1	.685	-.213	.531	.279	.304
Stretching 2	.356	.326	.591	.582	.400
Stretching 3	.787	-.082	.728	.484	.534
Stretching 4	.283	.442	.601	.645	.455
Stretching 5	.598	.105	.673	.535	.458
Embracing 1	-.186	.525	.192	.392	.170
Embracing 2	.614	.060	.657	.501	.433
Embracing 3	-.021	.697	.481	.682	.465
Embracing 4	.025	.670	.507	.688	.474
Embracing 5	.269	.242	.443	.436	.225
					Total
Sum of Squares (eigenvalues)	2.800	.490			3.290
Explained variance	34.995	6.126			41.121

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Loadings higher than .3 were given in bold

Table 54

Promax-Rotated Pattern and Structure Matrices for CEI-II in Study 3: Reduced Sets of Variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Stretching 1	.641	-.159	.541	.247	.307
Stretching 3	.709	.017	.719	.466	.518
Stretching 5	.606	.101	.670	.485	.455
Embracing 1	-.131	.464	.163	.381	.156
Embracing 3	-.006	.715	.447	.712	.506
Embracing 4	.099	.616	.489	.679	.466
					Total
Sum of Squares (eigenvalues)	1.950	.458			2.408
Explained variance	32.495	7.937			40.133

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

* Loadings higher than .3 were given in bold

As can be seen in Table 54, after elimination of problematic items, it was possible to obtain two clear structures with significantly loaded items. Factor 1 (Stretching) had eigenvalue of 1.950 and explained 32.495% of the total variance by itself. On the other, Factor 2 (Embracing) had eigenvalue of .458 and explained 7.937% of the total variance. In this respect, Factor 2 seemed to contribute to the explained total variance very little. However, since the three items loaded

significantly enough onto this factor and theoretically, it was logical, it was decided to retain these two factors.

In terms of the reliability of these two factors, Factor 1 had Cronbach's alpha value of .673 whereas Factor 2 had .588. Nunnally and Bernstein (1994) state that ideally alpha value should be between .70 and .95. In this sense, these findings indicated that these factors had low reliability values. Yet, George and Mallery (2003) point out that as a rule of thumb, alpha value between .90 and .80 is excellent, .80 and .70 is good, .70 and .60 is questionable, .60 and .50 is poor an below .50 is unacceptable (p. 231). In addition, as De Vellis (2003) argues, low reliability values could be resulted from low number of items, which was the case in this analysis. Therefore, based on these suggestions, both factors were retained as they were.

In terms of the validity of the two Factors, it could be said that since there were no items cross loaded onto both factors, these structures were unidimensional. In addition, seeing all the items highly loaded onto the factors, these structures had convergent validity. But, when the factor correlations, which are provided in Table 55, were investigated, it was found out that correlations provide evidence that the items in these two factors failed to discriminate with each other.

Table 55

Promax-Rotated Factor Correlation Matrix for CEI-II in Study 3

Factor	1	2
1	1.000	.633
2	.633	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

According to these findings, it could be stated that although the two factors had some problems with their construct validities resulting from discriminant validity problems, they could be used as suggested by their authors.

EFA with Achievement Goal Questionnaire-Revised in Study 3. First, the suitability of the data set was examined. According to the bivariate correlation matrix (Appendix 12), there were many correlations of .3 or higher. In addition, all the correlations were significant at .01 (2-tailed) level. Bartlett's Test of Sphericity, approx. chi-square = 2367,843, df = 15, $p < .001$, also indicated the significance of the correlations among variables. Moreover, Kaiser-Meyer-Olkin Measure of Sampling Adequacy value, .771, showed that the data was large enough to conduct

factor analysis. Therefore, at the next phase, exploratory factor analysis procedure with maximum likelihood extraction method was carried out.

Table 56, which includes the initial factor analysis results without rotation, shows that the two factors, the number of which was determined with the application of a priori criteria due to the theoretical considerations, initially explained 68.819% of the total variance. Factor 1 which had 2.927 eigenvalue and Factor 2 which had 1.202 eigenvalue contributed to this variance by 48.779% and 20.040% respectively.

Table 56

Unrotated Factor Matrix for AGQ-R in Study 3

	Factor		Communality
	1	2	
Mastery Avoidance 1	.376	.440	.335
Mastery Avoidance 2	.546	.577	.630
Mastery Avoidance 3	.337	.360	.244
Performance Avoidance 1	.792	-.092	.635
Performance Avoidance 2	.810	-.166	.684
Performance Avoidance 3	.825	-.212	.725
			Total
Sum of Squares (eigenvalues)	2.927	1.202	4.129
Explained variance	48.779	20.040	68.819

Extraction Method: Maximum Likelihood.

In terms of the communalities that each variable had, it could be said that each of them was above the required level. Thus, no action was taken to eliminate any of the variables. However, unrotated factors matrix was not enough to present clear factor structures. As Table 56 shows, Mastery Avoidance 1 and Mastery Avoidance 3 cross loaded onto the two factors. Therefore, it was decided to respecify the structures by applying Promax rotation.

Table 57

Promax-Rotated Pattern and Structure Matrices for AGQ-R in Study 3: Full Set of Variables

	Pattern Matrix*		Structure Matrix*		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Mastery Avoidance 1	-.020	.588	.281	.578	.335
Mastery Avoidance 2	.021	.783	.421	.794	.630
Mastery Avoidance 3	.010	.489	.260	.494	.244
Performance Avoidance 1	.755	.077	.794	.463	.635
Performance Avoidance 2	.829	-.003	.827	.421	.684
Performance Avoidance 3	.877	-.053	.850	.396	.725
					Total
Sum of Squares (eigenvalues)	2.516	.737			3.253

Explained variance	41.930	12.280	54.209

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.
* Loadings higher than .3 were given in bold

Table 57 presents the results of Promax-rotated factor analysis. Accordingly, it was found out that there was no need for item reduction since all the items loaded onto a factor significantly without any problem. Factor 1 (Performance Avoidance) and Factor 2 (Mastery Avoidance) together explained 54.209% of the total variance and they had 2.516 and .737 eigenvalues. Moreover, Factor 1 had .862 Cronbach's alpha value whereas Factor 2 had .646. In this respect, it may seem that factor 2 had a low reliability value which was questionable (George & Mallery, 2003) but there was no need to delete the factor due to its low number of items (DeVellis, 2003). Consequently, it could be stated that Mastery Avoidance and Performance Avoidance had internal consistency.

When construct validity of the two factors was examined, it could be said that since all the items loaded onto a single factor significantly high, these structures had convergent validity. However, when Table 58 was examined, it was seen that correlations between the two factors were relatively high. All in all, it was concluded that Mastery Avoidance and Performance Avoidance of AGQ-R could be used as suggested by their authors.

Table 58

Promax-Rotated Factor Correlation Matrix for AGQ-R in Study 3

Factor	1	2
1	1.000	.512
2	.512	1.000

Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.

Although the scales were examined in terms of their validities, only convergent validities and discriminant validities were discussed above. Another validity to be controlled was the face validity. In terms of face validity of the scales, some interviews were done with the volunteer participants. Especially, students were asked about the items in Scarcity scale which had significant problems. According to their comments and suggestions, necessary adjustments were made on the items of Scarcity.

In addition, experts who helped the translation back translation processes were counseled again. Accordingly, it was found out that there were some problems in terms of face validity of the whole scale. The first criticism was about font and

page sizes. Font style had been chosen as 'Times New Roman' in size 9 and page size had been adjusted to A5 (14.82 cm x 20.89 cm) in booklet format. Participants stated that it was lengthy to complete. Second, some participants stated that the first three items of DMC were a bit confusing and they had difficulty in comprehending them. Criticism made by the experts who indicated that numbering of items might have caused some problems because regardless of the part items belong to, all the items had been numbered in a continuous way starting from one. Another criticism by the experts was about the titles of the subscales which had been coded with letters (e.g.. Part A, B etc.) instead of the titles of the scales. Experts, also, pointed out that the formation of choices was confusing suggested using more reader friendly layout. Finally, experts demanded the section including demographic data which had been placed at the end to be placed at the beginning. In addition to above-mentioned issues, the item translations were reviewed by experts especially the ones which had low communalities or which had not been loaded onto any factors and some small adjustments were made.

Finally, for the next application a new version of the scale was created in A4 page size (20.99 cm x 29.70 cm) and font size was adjusted to 12 with a new layout which was more reader friendly. All the items in a section were numbered as a separate group and all section names were added. The part that included demographic data was placed at the beginning and one of DMC questionnaire which aimed at collecting data about the duration of DMC experience was deleted. Moreover, the number of choices in the first item of DMC dropped to three and simplified.

In addition, in terms of face validity in Study 2, there were also some problems. According to the interviews with the participants and experts, several adjustments were needed as in Study 1. Major criticism from the participants was about the length of the questionnaire. In addition, the feedback from the participants and experts revealed that Scarcity subscale had some problems regarding the translations of the items.

For the main application of Octalysis questionnaire, the number of items was reduced. Items which were not used in statistical analyses were eliminated. In addition, page layout, font style and size were also readjusted. Thus, it was possible to reduce the number of pages from 9 to 4. Another problem was about Scarcity

subscale which had been adapted from behavioral psychology. There were ongoing problems with Scarcity subscale both in Study 1 and Study 2. Therefore, before making new adjustments, some voluntary participants were asked to give their opinions about the five items in Scarcity. Participants stated that it was not easy to understand especially item numbered 2. Then, the participants were asked to provide an alternative statement. Accordingly, the original item which could be translated into English as “Even I am late, I enter my favorite class” was changed as “Even I have been written absent due to my late arrival, I enter to my favorite class”. Several minor adjustments were made in accordance with the feedbacks given by the two experts who are professors in English language teaching as a second language field.

Structural Equation Modelling

Having found out that all scales could be utilized as suggested by their authors in the previous studies, to examine the relationships between different constructs, more complicated multivariate techniques were needed. Among many alternatives, Structural Equation Modelling (SEM) comes to the fore since it has the capability to examine highly complex relationships (Hair et al., 2013, p. 542). SEM models can enable the examination of multiple interrelated dependence relationships. That is, as Ullman (2013) states, SEM is basically the combination of EFA and multiple regression analyses (p. 676). Besides, construct that is not observed directly can be represented within a SEM model and it is also possible to account for measurement error which is not possible with other multivariate techniques. Nevertheless, SEM is not free from some limitations. Ulmann (2007) states that SEM may be problematic in terms of both theoretical and practical respects. When the former is taken into consideration, it can be said that planning is highly essential for SEM analysis since it is used to test a theory. In this respect, it is crucial to have prior knowledge about the theory to be tested. The latter, SEM analysis can be very sensitive to some practical issues such as sample size and missing data, multivariate normality and outliers, linearity, absence of multicollinearity and singularity, and residuals (Ulmann, 2007; p. 683).

Traditionally, SEM models consist of a measurement model and a structural model (Hair et al., 2013; p. 551). However, it is important to note that all SEM models

should depend on strong theoretical bases because, as Hair et al. (2014) point out, SEM is particularly useful to test and confirm a theory. Once a model is developed on a sound theory, then it is possible to develop measurement model specification and structural model specification. According to Hair et al. (2013), there are six stages in SEM:

Stage 1: Defining individual constructs

Stage 2: Developing the overall measurement model

Stage 3: Designing a study to produce empirical results

Stage 4: Assessing the measurement model validity

Stage 5: Specifying the structural model

Stage 6: Assessing structural model validity

(Hair et al., 2013; p. 565).

In this study, how well the individual items in the scales measure the latent constructs and whether the constructs they formed were different from each other were tested. As suggested by Collier (2020), Confirmatory Factor Analysis was used for this procedure. Collier argues that EFA procedures, which were used in the previous studies, are useful for data reduction and can be seen as the first step to decide whether an item is the indicator of a construct.

Stage 1: Defining individual constructs. The hypothesized model which was a combination of eight different constructs is shown in Table 59. For space considerations, only the items used in CFA procedure were shown in the table. According to the model, these eight-factors are Epic Meaning and Calling, Development and Accomplishment, Empowerment of Creativity and Feedback, Ownership and Possession, Social Influence and Relatedness, Scarcity and impatience, Unpredictability and Curiosity, and Loss and Avoidance.

Table 59

Individual Items, Their Turkish Descriptions, Labels and Corresponding Constructs

Item	Scale type	Description	Label	Construct
DMC1	1-5 Likert Strongly Disagree – Strongly Agree	Geriye dönüp baktığımda bu süreçle ilgili güzel anılarım var.	DMC 1	Epic Meaning and Calling

DMC2	1-5 Likert Strongly Disagree – Strongly Agree	Bu süre içerisinde her zamankinden daha verimli bir şekilde çalışabilmiştim.	DMC 2	Epic Meaning and Calling
DMC5	1-5 Likert Strongly Disagree – Strongly Agree	Bu deneyim benim istediğimden de çok başarılı olmama yardımcı oldu.	DMC 5	Epic Meaning and Calling
DMC6	1-5 Likert Strongly Disagree – Strongly Agree	Sanırım bu süreç içerisinde bana özel bir şey oldu – çok müthiş bir zamandı.	DMC 6	Epic Meaning and Calling
DMC7	1-5 Likert Strongly Disagree – Strongly Agree	O zaman proje benim hayatımın merkezi haline geldi.	DMC7	Epic Meaning and Calling
DMC12	1-5 Likert Strongly Disagree – Strongly Agree	Gerçekten çok eğlenceli bir deneyimdi.	DMC 12	Epic Meaning and Calling
AGQR5	1-5 Likert Strongly Disagree – Strongly Agree	Amacım yapabileceğimden daha azını öğrenmekten kaçınmaktır.	Mastery Avoidance 1	Loss and Avoidance
AGQR9	1-5 Likert Strongly Disagree – Strongly Agree	Amacım öğrenilmesi gerekenden daha az öğrenmekten kaçınmaktır.	Mastery Avoidance 2	Loss and Avoidance
AGQR11	1-5 Likert Strongly Disagree – Strongly Agree	Ders konularını yarım yamalak anlamaktan kaçınmaya gayret ediyorum.	Mastery Avoidance 3	Loss and Avoidance
AGQR6	1-5 Likert Strongly Disagree – Strongly Agree	Amacım diğer öğrencilere kıyasla zayıf performans göstermekten kaçınmaktır.	Performance Avoidance 1	Loss and Avoidance
AGQR10	1-5 Likert Strongly Disagree – Strongly Agree	Diğer öğrencilerden daha kötü performans göstermekten kaçınmaya gayret ediyorum.	Performance Avoidance 2	Loss and Avoidance
AGQR12	1-5 Likert Strongly Disagree – Strongly Agree	Amacım diğer öğrencilerden daha başarısız olmaktan kaçınmaktır.	Performance Avoidance 3	Loss and Avoidance
BPNSF1	1-5 Likert Not True At All – Completely True	Üstlendiğim işlerde seçim şansım olduğunu ve özgür olduğumu hissediyorum.	Autonomy Satisfaction 1	Empowerment of Creativity and Feedback
BPNSF3	1-5 Likert Not True At All – Completely True	Önemsediğim insanların da beni önemsediklerini hissediyorum.	Relatedness Satisfaction 1	Social Influence and Relatedness
BPNSF5	1-5 Likert Not True At All – Completely True	İşleri iyi yapabileceğim konusunda kendimden eminim.	Competence Satisfaction1	Development and Accomplishment
BPNSF7	1-5 Likert Not True At All – Completely True	Verdiğim kararların gerçekten istediğim şeyleri yansıttığını hissediyorum.	Autonomy Satisfaction 2	Empowerment of Creativity and Feedback
BPNSF9	1-5 Likert Not True At All – Completely True	Benim önemsediğim ve beni önemseyen insanlarla bağlantıda olduğumu hissediyorum.	Relatedness Satisfaction 2	Social Influence and Relatedness
BPNSF11	1-5 Likert Not True At All – Completely True	Yaptığım şeyde yetenekli olduğumu hissediyorum.	Competence Satisfaction 2	Development and Accomplishment
BPNSF13	1-5 Likert Not True At All – Completely True	Yaptığım seçimlerin beni ifade ettiğini hissediyorum.	Autonomy Satisfaction 3	Empowerment of Creativity and Feedback
BPNSF15	1-5 Likert Not True At All – Completely True	Benim için önemli olan insanlara karşı kendimi onlara yakın ve onlarla bağlantılı hissediyorum.	Relatedness Satisfaction 3	Social Influence and Relatedness

BPNSF17	1-5 Likert Not True At All – Completely True	Amaçlarımı gerçekleştirebileceğim konusunda kendimden eminim.	Competence Satisfaction 3	Development and Accomplishment
BPNSF19	1-5 Likert Not True At All – Completely True	Gerçekten ilgim olan şeyleri yapmakta olduğumu hissediyorum.	Autonomy Satisfaction 4	Empowerment of Creativity and Feedback
BPNSF21	1-5 Likert Not True At All – Completely True	Beraber vakit geçirdiğim insanlara karşı samimi duygular yaşıyorum.	Relatedness Satisfaction 4	Social Influence and Relatedness
BPNSF23	1-5 Likert Not True At All – Completely True	Zor görevleri başarı ile tamamlayabileceğimi hissediyorum.	Competence Satisfaction 4	Development and Accomplishment
CEI1	1-5 Likert Very Slightly or Not at all -Extremely	Yeni bir şeyle karşılaştığım durumlarda olabildiğince çok bilgi edinmenin yollarını ararım.	Stretching 1	Unpredictability and Curiosity
CEI3	1-5 Likert Very Slightly or Not at all - Extremely	Karmaşık ya da zorlu şeyleri yapmada çok iyiyimdir.	Stretching 2	Unpredictability and Curiosity
CEI5	1-5 Likert Very Slightly or Not at all - Extremely	Zorlu şeyleri birer gelişme ve öğrenme fırsatı olarak görürüm.	Stretching 3	Unpredictability and Curiosity
CEI7	1-5 Likert Very Slightly or Not at all - Extremely	Her zaman kedimle ilgili düşüncelerimi ya da dünya görüşümü zorlayıcı deneyimleri ararım.	Stretching 4	Unpredictability and Curiosity
CEI9	1-5 Likert Very Slightly or Not at all - Extremely	Sık sık beni zorlayan ve kişisel gelişimime katkısı olacak fırsatları ararım.	Stretching 5	Unpredictability and Curiosity
POQ1	1-5 Likert Strongly Disagree – Strongly Agree	Hazırlık İngilizce programının benim programım olduğunu hissediyorum	Psychological Ownership 1	Ownership and Possession
POQ2	1-5 Likert Strongly Disagree – Strongly Agree	Hazırlık İngilizce programına karşı “kişisel sahiplik” duygusuna sahibim	Psychological Ownership 2	Ownership and Possession
POQ3	1-5 Likert Strongly Disagree – Strongly Agree	Hazırlık İngilizce programının benim programım olduğunu çok zor hissediyorum.	Psychological Ownership 3	Ownership and Possession
POQ5	1-5 Likert Strongly Disagree – Strongly Agree	Hazırlık İngilizce programı tam benim programım.	Psychological Ownership 5	Ownership and Possession
STPS1	1-5 Likert Strongly Disagree – Strongly Agree	Nadir karşılaştığım öğrenme materyallerini ve aktivitelerini daha değerli bulurum.	Scarcity 1	Scarcity and impatience
STPS3	1-5 Likert Strongly Disagree – Strongly Agree	Nadir yapılan aktivitelerin katılımcılarından birisi olduğumda kendimi iyi hissederim.	Scarcity 3	Scarcity and impatience
STPS4	1-5 Likert Strongly Disagree – Strongly Agree	Sevdiğim bir materyal ya da aktivite bitmek üzereyse üzülürüm.	Scarcity 4	Scarcity and impatience
STPS5	1-5 Likert Strongly Disagree – Strongly Agree	Sınıfta elde edilmesi zor olan ödüllerin özel bir değeri vardır.	Scarcity 5	Scarcity and impatience

As can be seen from the above Table 59, the model which was aimed to be tested in this study had borrowed scales from different studies. Hair et al. (2013) indicate that in such cases, a pretest should be carried out with the participants who

have similar characteristics with the population who take part in the main study and a screening procedure needs to be carried out to see whether it is suitable to use those items in the study (p. 576). As previously addressed, two separate EFA procedures had been carried out (Piloting 1 and Piloting 2) with the participation of students who had been in the same language preparatory programs in three different universities. In addition, another EFA procedure was run at the final stage and then, a confirmatory factor analysis was performed with the data collected at the final stage of the study.

Based on the theory proposed by Chou (2015), the current study was designed focusing on the eight core drives underlying students' behaviors during their second language learning experiences and eventually effect their long-term motivation. The working definitions of the eight constructs were as follows:

- Epic Meaning and Calling. Students' desire to reach a meaningful and great goal throughout their language learning experiences.
- Development and Accomplishment. The extend students are motivated to complete a challenge and accomplish an objective. When they feel they develop and achieve something in terms of their language proficiency, they tend to go on striving to learn.
- Empowerment of Creativity and Feedback. The extend students want to be autonomous and have their own learning responsibilities and intrinsically motivated to learn. If they enjoy learning a second language just for its own sake, they feel good and take action easily.
- Ownership and Possession. Students perform better in their long-term language learning experiences if they feel the ownership of the experience.
- Social Influence and Relatedness. Students perform better when they are a part of a group because they have the desire to connect and compare with one another.
- Scarcity and impatience. When materials which are scarce and which attracts students and make them act immediately tend to increase their motivation.

- Unpredictability and Curiosity. It is important to foster curiosity so that individuals who recognize, pursuit, and try to integrate novel and challenging experiences and information in their language learning experiences be more motivated.
- Loss and Avoidance. Sometimes students are motivated only for not to lose. They may not have a meaningful and bigger goal but may want to preserve their current position and do the minimum.

When the scale which had been formed by bringing together the above-mentioned subscales from different sources was examined in terms of its face validity, it can be said that face validity was established. In addition, conceptual definitions and translations of these definitions match with the item wordings. Four different experts examined all the items and their translations and they had no difficulty in confirming the scale.

Stage 2: Developing the overall measurement model. Having specified the constructs, the following step included the specification of measurement model. A visual diagram of measurement model is provided in Figure 13. The model consists of eight latent constructs, which were shown with oval shapes, measured by 37 indicator variables, which were given with boxes with the labels as they were used in the questionnaire. One of these latent constructs which was Loss and Avoidance two sub-constructs (CD7a and CD7b) that had items measuring Mastery Avoidance and Performance Avoidance dimensions of the model. One-headed arrows indicated a casual path from the construct to the measured indicator. In other words, all these measures were considered to be reflected and the causality was from the latent construct to the measured items. Items measuring each single construct were supposed to be conceptually similar and empirically acted together.

Since there were no reasons to think that all these latent constructs were independent from each other, they all correlated with each other with two headed arrows between constructs. As seen from the Figure 13, all measured variables loaded only on one construct and had error terms, which were shown with the term e and attached to each measure variable, did not relate to a different variable. Among these eight latent constructs, Epic Meaning and Calling was indicated by six items whereas five constructs (Scarcity and Impatience, Development and

Accomplishment, Empowerment of Creativity and Feedback, Ownership and Possession, Social Influence and Relatedness) were indicated by four measured items. Unpredictability and Curiosity was measured with five indicators and finally, six items which were grouped under two sub-constructs indicated the last construct (Loss and Avoidance). Each construct was identified and the model had more degrees of freedom than the paths to be estimated and thus, it was possible to state that the model was overidentified (Hair et al., 2013; p.628).

Stage 3: Designing a study to produce empirical results. According to Hair et al. (2013), the development of the overall measurement model in the second step is followed by the design of a study and collection of data to test the measurement model. In this respect, this study was designed to obtain empirical data to test the measurement model. Initially, data were obtained from 1650 participants from six universities across Turkey (Muğla Sıtkı Koçman University, Çanakkale Onsekiz Mart University, Erzurum Atatürk University, Pamukkale University, Burdur Mehmet Akif Ersoy University, and Izmir Institute of Technology). After the data screening procedures had been applied in EFA stages of current study, this number dropped to 1150. Since the data screening procedures had been carried out in the previous analyses, no further data screening procedure was applied at this stage. as they were suggested by their authors. In addition, relatively large sample size made it possible to use the maximum likelihood estimation method.

After the development of the measurement model at the previous stage, the specification and estimation of the model were carried out. Estimated parameters and degrees of freedom are shown in Table 60 which includes CFA results.

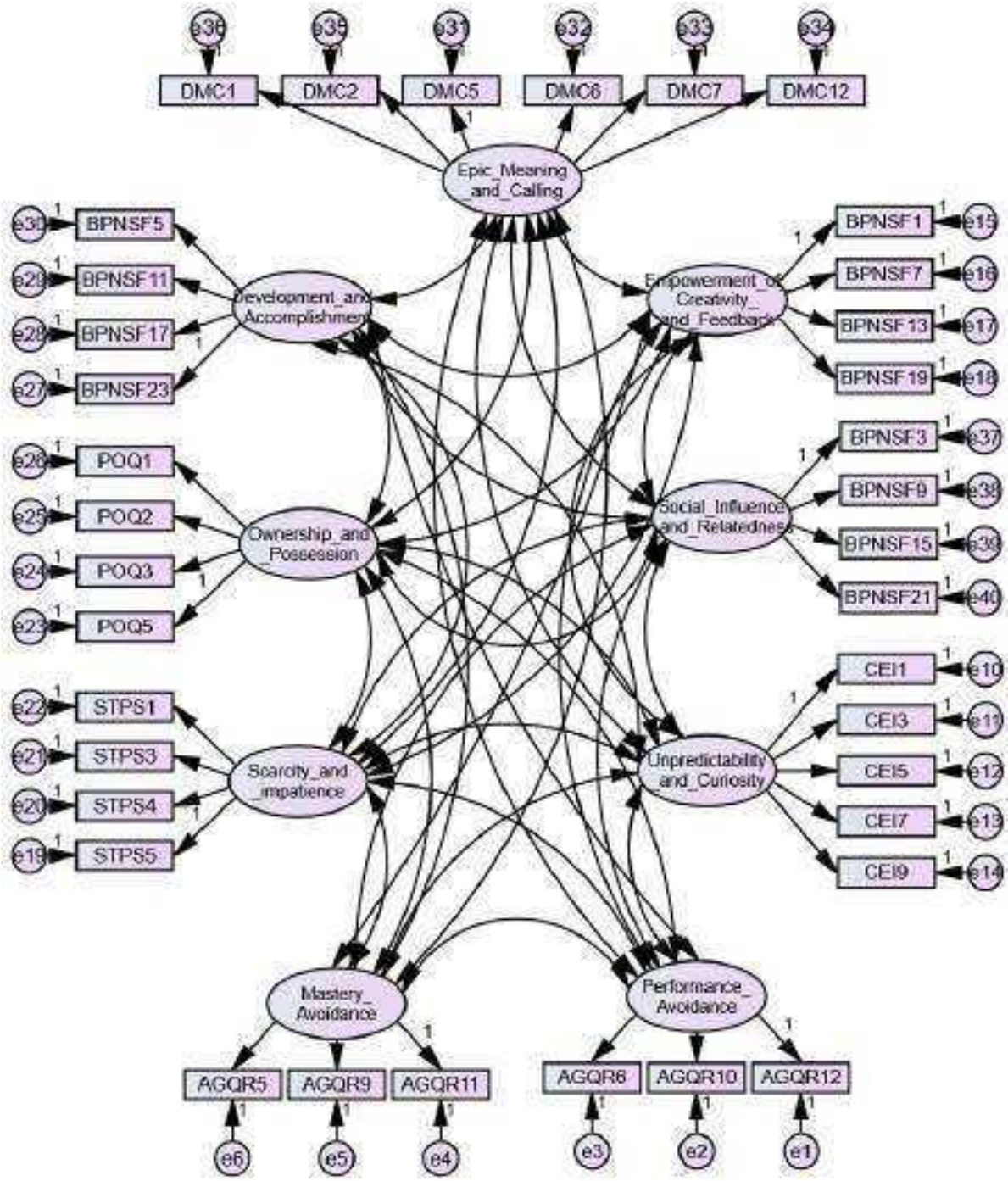


Figure 13. Measurement Model for Octalysis Theory

Table 60

Estimated Parameters in Octalysis CFA Model

	Variable Label*								
Variable Name	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8a	CD8b
DMC1	0	0	0	0	0	0	0	0	0
DMC2	1	0	0	0	0	0	0	0	0
DMC5	2	0	0	0	0	0	0	0	0
DMC6	3	0	0	0	0	0	0	0	0
DMC6	4	0	0	0	0	0	0	0	0
DMC12	5	0	0	0	0	0	0	0	0
BPNSF1	0	0	0	0	0	0	0	0	0
BPNSF7	0	6	0	0	0	0	0	0	0
BPNSF13	0	7	0	0	0	0	0	0	0
BPNSF19	0	8	0	0	0	0	0	0	0
BPNSF5	0	0	0	0	0	0	0	0	0
BPNSF11	0	0	9	0	0	0	0	0	0
BPNSF17	0	0	10	0	0	0	0	0	0
BPNSF23	0	0	11	0	0	0	0	0	0
BPNSF3	0	0	0	0	0	0	0	0	0
BPNSF9	0	0	0	12	0	0	0	0	0
BPNSF15	0	0	0	13	0	0	0	0	0
BPNSF21	0	0	0	14	0	0	0	0	0
POQ1	0	0	0	0	0	0	0	0	0
POQ2	0	0	0	0	15	0	0	0	0
POQ3	0	0	0	0	16	0	0	0	0
POQ5	0	0	0	0	17	0	0	0	0
CEI1	0	0	0	0	0	0	0	0	0
CEI3	0	0	0	0	0	18	0	0	0
CEI5	0	0	0	0	0	19	0	0	0
CEI7	0	0	0	0	0	20	0	0	0
CEI9	0	0	0	0	0	21	0	0	0
STPS1	0	0	0	0	0	0	0	0	0
STPS3	0	0	0	0	0	0	22	0	0
STPS4	0	0	0	0	0	0	23	0	0
STPS5	0	0	0	0	0	0	24	0	0
AGQR5	0	0	0	0	0	0	0	0	0
AGQR9	0	0	0	0	0	0	0	25	0
AGQR11	0	0	0	0	0	0	0	26	0
AGQR6	0	0	0	0	0	0	0	0	0
AGQR10	0	0	0	0	0	0	0	0	27
AGQR12	0	0	0	0	0	0	0	0	28
Construct variances and covariances									
CD1	29								
CD2	30	36							

CD3	31	37	44						
CD4	32	38	45	51					
CD5	33	39	46	52	57				
CD6	34	40	47	53	58	62			
CD7	33	41	48	54	59	63	66		
CD8a	34	42	49	55	60	64	67	69	
CD8b	35	43	50	56	61	65	68	70	71
Error terms for indicators = 39									
Total number of estimated parameters: 28 + 45 + 39 = 112									

* CD1: Epic Meaning and Calling; CD2: Development and Accomplishment; CD3: Empowerment of Creativity and Feedback; CD4: Ownership and Possession; CD5: Social Influence and Relatedness; CD6: Scarcity and Impatience; CD7: Unpredictability and Curiosity; CD8a: Loss and Avoidance (Mastery Avoidance), CD8b: Loss and Avoidance (Performance Avoidance)

As can be seen in Table 60, the model had 96 parameters. Of these 96 free parameters, 26 were factor loadings, 36 were factor variance and covariance, and 34 were error terms. The unique variance and covariance terms total was: (39 x 40)/2= 780. Since 780 was greater than 112, it was possible to say that the model was identified and included more degrees of freedom than free parameters.

Stage 4: Assessing measurement model validity. As Ullman (2013) points out after the model has been formed, it is time to decide whether this model is a good one. And one of the ways to make a decision is to check the conformity between sample covariance matrix and population covariance matrix which is estimated by the model (p.720). The perfect fit means that the two matrices are the same. Since perfect fit would not be possible, it could be said that the more similar the two matrices, the better fit exists. Therefore, to confirm model validity, goodness-of-fit is to be sought. Stage 4 included the comparison of the model theoretically developed to the one in reality. According to Hair et al. (2013), a goodness-of-fit (GOF) at or above the suggested level and acceptable construct validity are the two critical factors affecting the model validity (p.576). GOF measures are grouped into three categories: absolute, incremental and parsimony fit indices (Hair et al., 2013; p. 576). Values of GOF indices at acceptable levels indicate the existence of good model whereas inconsistency among several indices requires re-examination of the model (Ullman, 2013; p. 720). However, there is not a consensus among researchers about which indices to report. Bentler and Bonett (1980) state that mostly used fit indices such as CFI, TLI, NFI, and IFI would be enough to report. Hu and Bentler (1999) suggest reporting SMRM index and a comparative fit index. Hair

et al. (2013) suggest using chi-square (x^2) value and degrees of freedom, one incremental index and one absolute index while reporting CFA results. And in the same way, Kline (2011) points out that while reporting research results, degrees of freedom and chi-square statistics with the p-value, root mean square error of approximation (RMSEA) and its ninety percent confidence interval, comparative fit index (CFI), and standardized root mean square residual (SRMR) should be included.

In this respect, to check the overall fit, some fit indices suggested by Hair et al. (2013; p. 630) were provided in Table 61. Hair et al., (2013) suggest that together with x^2, at least one absolute fit index and one incremental fit index should be reported (p.630) in studies. When Table 61 is checked, it is seen that the x^2 for the current study was 1243.744 with significant p value (p = .000). Normally, in many other multivariate statistical tests, significant p-values show that assumptions are met. Yet, Hair et al. (2013) state that different from other statistical tests, x^2 value should be insignificant because significant x^2 shows that observed and estimated covariance matrices are statistically different from each other whereas insignificant x^2 means equal covariance matrices. That is, x^2 value should not be lower than .05 (p. 577). Hair et al. (2013) also argue that when the sampling size exceeds 250 and the number of observed variables is over 30, significant p-values are expected (p. 584).

Table 61

Octalysis CFA Goodness-of-Fit Statistics

Chi-square (x^2)
Chi-square (CMIN) = 1243.744 (p = .000)
Degrees of freedom (DF) = 593
CMIN/DF = 2.097
Absolute Fit Measures
Goodness-of-fit index (GFI) = .945
Root mean square error of approximation (RMSEA) = .031
90 percent confidence interval for RMSEA = .028; .033
Root mean square residual (RMR) = .039
Squared root mean square residual (SRMR) = .0363
Incremental Fit Indices
Normed fit index (NFI) = .908
Comparative fit index (CFI) = .949
Relative fit index (RFI) = .896
Parsimony Fit Indices
Adjusted goodness-of-fit index (AGFI) = .935
Parsimony normed fit index (PNFI) = .808

Moreover, Hair et al. (2013), Ullman (2013), Gulliksen and Tukey (1958), Collier (2020) and Jöreskog and Sörbom (1969) argue that there is a close relationship with x^2 and the sample size. Generally, as the sample size increases, the probability of obtaining an insignificant x^2 ($x^2 > .05$) decreases. In other words, large sample sizes yield significant x^2. In addition, Bollen and Long (1993) state that even the assumptions are met, it is questionable to use only x^2 test to determine the fit of the model because as Kline (2011) says, any decision depending only on x^2 results is not enough to reject or keep the model. Thus, Collier (2020) suggests using relative chi-square values (CMIN/DF) to obtain results which are less dependent on sample size (p. 66). Carmines and McIver (1981) point out that CMIN/DF values between 2 to 1 or 3 to 1 ratio show adequate fit between covariance matrices (p.80). Byrne(1989) says that CMIN/DF value should not exceed 2 for good fit (p. 55). And Kline (2011) argues that CMIN/DF value under 3 is the indicator of acceptable fit. The findings of this study showed that when x^2 was divided into degrees of freedom, the CMIN/DF value was 2.097, which was below the suggested level and indicated a good fit.

However, Kline (2011) says that any decision depending only on x^2 results is not enough to reject or keep the model. Thus, chi-square statistics alone is not enough to determine the fitness of the model. Ulmann (2007) proposes reporting other indices from other groups (p.720). An absolute fit index, root mean square error of approximation (RMSEA), is one of the commonly used fit indices. Hair et al. (2013) say that RMSEA indicates “how well a model fits a population” (p. 579) and Kline (2011) points out that the closer RMSEA to zero, the better fit it indicates (p. 273). Browne and Cudeck (1993) state that a RMSEA value of .05 or less indicates a close fit and .08 or less is also acceptable. But they add that RMSEA value should not exceed .1. Similarly, researchers such as Brown (2006), Jöreskog and Sörbom (1969), Sümer (2000) and Collier (2020) argue that RMSEA below .05 would point out a perfect model fit. However, there are other researchers who accept looser cut-off point. Thompson (2004) states that values up to .06 can be accepted as the indicators of good fit whereas Steiger (2007) says this limit can be .07. Sümer (2000) and Hooper, Coughlan and Mullen (2008) maintain looser point of view and state that RMSEA values up to .08 can signal good fit while Ullman (2013) says that values larger than .1 show weak fit between the matrices. However, Hair et al.

(2013) argue that giving a certain cut-off point for RMSEA is not advisable. Instead, RMSEA values should be reported regarding their ranges at certain percentage of confidence intervals (p. 579). When Table 61 is examined, it could be seen that a RMSEA value of .031 (between .028 and .033 with 90% confidence interval) was obtained. And by taking the above-mentioned suggestions by the researchers, it would not be wrong to state that the model tested in the scope of this study had a good fit. In addition, standardized root mean square residual (SRMR), which is advised to be reported when sample size exceeds 500 (Kline, 1998), was found .036, which was below the cut-off value .05 and signals the good fit between matrices (Hair et al., 2013; p.630).

Another frequently used index is the Comparative fit index (CFI), which is one of the incremental fit indices whose values range from 0 to 1. CFI evaluates how well the model fits to the reality. The higher the value in this range, the better fit it indicates (Hair et al., 2013). As Collier (2020) points out that sample size does not have an effect on CFI, thus it is recommended to be included into reports. According to Bentler and Bonett (1980) and Hair et al. (2013), for acceptable model fit, CFI values need to be at or above .90. Hu and Bentler (1999), however, state that to obtain good fit, CFI should be above .95. Findings of this study yielded a CFI value of .949, which was above the suggested levels and indicates the existence of good fit between the covariance matrix produced in this study and observed covariance matrix. All in all, it can be stated that although it is possible to report many more indices, the above-mentioned indices provide enough evidence that support good model fit.

Having found out that the model has a fit to the data, the second step of deciding the validity of the measurement model includes the examination of construct validity which shows how accurate the measured items in the model represents the latent constructs. It consists of four components which are: convergent validity, discriminant validity, nomological validity and face validity. (Hair et al., 2013; p. 618).

When items in a construct converge and are indicators of a specific construct, it is called as convergent validity. Hair et al. (2013) point out that factor loadings of items need to be checked to assess the convergent validity. It can be said that high factor loadings are the indicators of the high convergent validity. At the minimum, all

loadings of all items should be significant to be able to talk about the existence of convergent validity. Table 62 shows the unstandardized factor loadings, standard errors, t-values (critical ratios) and p values for the tested model. It can be seen that all standardized factor loadings were significant at $p < .001$ level, which indicates the existence of convergent validity. However, as Hair et al. (2013) state, unstandardized factor loading can provide diagnostic information. To be able to calculate discriminant validity and reliability, standardized loadings should be examined (p.632).

Table 62

Unstandardized Estimates, Standard Error, t-Values and p-Values for Octalysis CFA

Item	Construct	Estimate	S.E.	T Vlaue	P Value
DMC5	CD1: Epic Meaning and Calling	1	_a	_a	_a
DMC6	CD1: Epic Meaning and Calling	1.231	.074	16.556	***
DMC7	CD1: Epic Meaning and Calling	.793	.071	11.246	***
DMC12	CD1: Epic Meaning and Calling	1.138	.074	15.318	***
DMC2	CD1: Epic Meaning and Calling	.867	.062	14.056	***
DMC1	CD1: Epic Meaning and Calling	1.019	.063	16.214	***
BPNSF23	CD2: Development and Accomplishment	1	_a	_a	_a
BPNSF17	CD2: Development and Accomplishment	1.065	.053	20.11	***
BPNSF11	CD2: Development and Accomplishment	.988	.051	19.303	***
BPNSF5	CD2: Development and Accomplishment	.987	.051	19.329	***
BPNSF1	CD3: Empowerment of Creativity and Feedback	1	_a	_a	_a
BPNSF7	CD3: Empowerment of Creativity and Feedback	1.081	.072	15.113	***
BPNSF13	CD3: Empowerment of Creativity and Feedback	1.251	.079	15.866	***
BPNSF19	CD3: Empowerment of Creativity and Feedback	1.147	.079	14.48	***
POQ5	CD4: Ownership and Possession	1	_a	_a	_a
POQ3	CD4: Ownership and Possession	.886	.033	27.086	***
POQ2	CD4: Ownership and Possession	.893	.028	32.176	***
POQ1	CD4: Ownership and Possession	1.055	.028	37.83	***
BPNSF3	CD5: Social Influence and Relatedness	1	_a	_a	_a
BPNSF9	CD5: Social Influence and Relatedness	.969	.084	11.473	***
BPNSF15	CD5: Social Influence and Relatedness	1.188	.091	13.117	***
BPNSF21	CD5: Social Influence and Relatedness	.816	.069	11.753	***
STPS5	CD6: Scarcity and Impatience	1	_a	_a	_a
STPS4	CD6: Scarcity and Impatience	.623	.068	9.172	***
STPS3	CD6: Scarcity and Impatience	.833	.074	11.199	***
STPS1	CD6: Scarcity and Impatience	.826	.073	11.3	***
CEI1	CD7: Unpredictability and Curiosity	1	_a	_a	_a
CEI3	CD7: Unpredictability and Curiosity	1.299	.089	14.561	***
CEI5	CD7: Unpredictability and Curiosity	1.462	.094	15.593	***
CEI7	CD7: Unpredictability and Curiosity	1.367	.098	13.932	***
CEI9	CD7: Unpredictability and Curiosity	1.499	.098	15.235	***
AGQR11	CD8a: Mastery Avoidance	1	_a	_a	_a
AGQR9	CD8a: Mastery Avoidance	1,632	.119	13.733	***
AGQR5	CD8a: Mastery Avoidance	1,213	.094	12.848	***
AGQR12	CD8b: Performance Avoidance	1	_a	_a	_a
AGQR10	CD8b: Performance Avoidance	0,906	.03	29.828	***
AGQR6	CD8b: Performance Avoidance	0,913	.031	29.017	***

_a not estimated as loadings were fixed to 1.0; *** p < .001

Table 63 provides information for standardized factor loadings, average variance extracted, and construct reliability. Hair et al. (2013) state that loadings should be .5 or higher at least but for ideal results .7 or higher factor loadings are better. Besides, they state that “at minimum all factor loadings should be statistically significant” (p.618).

Table 63

Standardized Factor Loadings, Average Variance Extracted Values and Construct Reliability Values for Each Structure of Octalysis Model

Variable Name	Variable Label*								
	CD1	CD2	CD3	CD4	CD5	CD6	CD7a	CD7b	CD8
DMC1	.66								
DMC2	.54								
DMC5	.60								
DMC6	.68								
DMC7	.41								
DMC12	.60								
BPNSF5		.66							
BPNSF11		.66							
BPNSF17		.69							
BPNSF23		.71							
BPNSF1			.54						
BPNSF7			.66						
BPNSF13			.73						
BPNSF19			.61						
POQ1				.92					
POQ2				.80					
POQ3				.72					
POQ5				.83					
BPNSF3					.53				
BPNSF9					.50				
BPNSF15					.73				
BPNSF21					.52				
STPS1						.55			
STPS3						.54			
STPS4						.39			
STPS5						.54			
CEI1							.53		
CEI3							.63		
CEI5							.71		
CEI7							.58		

CEI9							.68		
AGQR5								.58	
AGQR9								.77	
AGQR11								.53	
AGQR6									.80
AGQR10									.83
AGQR12									.84
Average Variance Extracted	.35	.46	.41	.68	.33	.26	.40	.40	.68
Composite Reliability	.77	.78	.73	.90	.66	.58	.76	.66	.86

* CD1: Epic Meaning and Calling; CD2: Development and Accomplishment; CD3: Empowerment of Creativity and Feedback; CD4: Ownership and Possession; CD5: Social Influence and Relatedness; CD6: Scarcity and Impatience; CD7: Unpredictability and Curiosity; CD8a: Loss and Avoidance (Mastery Avoidance), CD8b: Loss and Avoidance (Performance Avoidance)

As can be seen from Table 63, except for the items DMC7 and STPS4, all the other items had factor loadings ranging between .92 and .50, which are above the minimum level suggested by Hair et al. and show that items adequately converge (2013). However, DMC7 and STPS4 had factor loadings .41 and .39 respectively. Researchers such as Hair et al. (2013) and Collier (2020) indicate that as a rule of thumb, item loadings need to be higher than .5. However, Garson (2010) and Stevens (1992) state that loadings as low as .4 can be kept as weak loadings. Moreover, Hair et al. (2013) indicate that any decisions regarding the retention of items in a structure should be taken based on the theory. In theory, Octalysis has four levels and at each level it requires different combinations of proposed CDs. Therefore, when Octalysis theory is taken into consideration, it is seen that some CDs such as Scarcity work better at first level whereas they should not be preferred for long term goals because they cause problems in terms of sustaining long term motivation. In addition, Chu (2015) points out that Scarcity is expected to have the least impact on long-term motivation. Depending on the above-mentioned references and considering the theory, it was decided to retain those two items (STPS2 and STPS4) although they have low loadings.

In this respect, to determine the convergent validity of the structures, the next criterion is average variance extracted (AVE) which is obtained by adding all squared standardized factor loadings and dividing them by the number of items (Fornell & Larcker, 1981; Hair et al., 2013). Collier (2020) states that AVE of .5 or higher indicates that items converge adequately. Values below .5 show that variance resulted from the measurement error is larger than the variance explained

by the construct which makes the validity of the construct questionable. When Table 63 is examined, it is seen that Development and Accomplishment, Ownership and Possession and Performance Avoidance dimensions of the model have AVE values at or above .5. However, the other dimensions of Octalysis model had AVEs lower than the suggested level, which indicated that structures failed to meet required criteria for adequate convergent validity. But Fornell and Larcker (1981) indicate that AVE is rather strict to determine the convergent validity of the structures, they add that “the researcher may conclude that the convergent validity of the construct is adequate even though more than 50% of the variance is due to error” (p.46), which means that it is possible to use structures with AVE less than .5. With a similar point of view Huang, Wang, Wu and Wang (2013) argue that although AVE of .5 is desired level, it is possible to accept AVE of .4 (p.219). Steinmetz (2016) point out that CR alone is a good indicator of the existence of convergent validity. Similarly, Fornell and Larcker (1981; p. 9), and Lam (2012) also argue that when AVE is less than .5 but composite reliability is .6 or higher, then it is possible to assume that convergent reliability is adequate. Therefore, although average variance extracted estimates fell below the desired level for adequate convergence, before reaching a conclusion, first CR values for each construct were investigated.

Composite reliability (CR) is the proportion of true variance to total variance. Hair et al. (2013) mention that reliability between .6 and .7 is adequate whereas values of .7 and above support good reliability. (p. 619). The higher the reliability, the more consistent constructs exist. Fornell and Larcker (1981) argue that instead of making judgements based on AVE, composite reliability can be alone used as an alternative. Similarly, Gaskin (2018) finds AVE of .5 too strict as a cut-off point and adds that values below .5 could be kept and CR values could be used to determine the validity of the constructs. Kline (2011) and Malhotra and Dash (2011), who have the same point of view, find AVE more conservative than CR and state that researchers can determine whether the convergent validity of the construct is adequate by taking CR values into consideration alone (p.702). In this respect, When Table 63 is investigated, it could be seen that CR values for each construct range from .9 to .6, which indicate that it was possible to assume that convergent validity was adequate.

Another part of construct validity to be controlled before further analyses is the discriminant validity, which is calculated with the shared variance between constructs (Fornell & Larcker, 1981). Once the correlation between constructs is obtained through a correlation matrix, they need to be squared and the result is expected to be greater that the AVE of each construct to assume that discriminant validity exists (Collier, 2020; p. 83). However, Collier (2020) criticizes assessing discriminant validity in the way proposed by Fornell and Larcker (1981). Henseler, Ringle and Sarstedt (2015) with a similar point of view indicate that traditional way of discriminant validity, which is Fornell-Larcker criterion, is unable to detect the lack of discriminant validity as a reliable tool and they recommend using Heterotrait-Monotrait Ratio of Correlations (HTMT) (p. 115). What is examined with HTMT method is "the correlations of indicators across constructs to the correlations of indicators within a construct" (Collier, 2020; p. 85). According to Henseler et al. (2015), the acceptable level of HTMT is .9. Kline (2011) points out that HTMT value above .85 is the indicator of discriminant validity problems.

Table 64

Heterotrait-Monotrait Ratio of Correlations

	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8a	CD8b
CD1: Epic Meaning and Calling	1								
CD3: Empowerment of Creativity and Feedback	.389	1							
CD2: Development and Accomplishment	.413	.708	1						
CD5: Social Influence and Relatedness	.304	.423	.403	1					
CD4: Ownership and Possession	.312	.398	.239	.434	1				
CD7: Unpredictability and Curiosity	.482	.474	.636	.212	.167	1			
CD6: Scarcity and Impatience	.402	.297	.326	.305	.294	.423	1		
CD8a: Mastery Avoidance	.357	.309	.298	.271	.270	.355	.385	1	
CD8b: Performance Avoidance	.150	.067	.065	.111	.141	.081	.315	.530	1

When the HTMT values provided in Table 64 are investigated, it could be seen that HTMT values in the measurement model range between .708 and .067, none of which were below the suggested threshold of .85. These findings indicated that discriminant validity in the measurement model was established.

In addition to examining the measurement model in terms of goodness of fitness and construct validity, Hair et al. (2013) advise that other diagnostic cues

such as standardized residuals, modification indices and specification search should be investigated to see whether it is necessary to modify the measurement model (p. 620).

The difference between observed and estimated covariance terms is called as residuals which can both have positive and negative values. As the fit gets better between covariance terms, residuals are expected to get smaller. Standardized residuals can be calculated by dividing residuals by the standard error of residuals. These values can be utilized to measure the strength of the difference between observed and expected terms (Glen, 2013). According to Hair et al. (2013), when standardized residuals are less than ± 2.5, it means there is not a problem. Standardized residuals between ± 2.5 and ± 4.0 should be examined but if there are no other problems with the items, they could be kept. However, values above ± 4.0 are unacceptable and require item be deleted (p. 621).

Table 65

Model Diagnostic Tools for the Octalysis Measurement Model: Standardized Residuals

Standardized residuals (residuals greater than ± 2.5)			
Negative standardized residuals			
CEI9	and	BPNSF9	-2.613
CEI7	and	BPNSF9	-3.579
DMC12	and	DMC2	-2.512
CEI9	and	DMC2	-2.718
DMC7	and	DMC12	-2.584
BPNSF13	and	STPS4	-3.150
CEI7	and	STPS5	-3.304
CEI3	and	STPS5	-2.507
Positive standardized residuals			
STPS3	and	BPNFS21	2.622
CEI1	and	BPNFS21	2.784
CEI1	and	DMC1	3.023
DMC2	and	DMC7	2.778
BPNFS19	and	DMC12	2.794
BPNFS1	and	DMC12	3.053
BPNFS19	and	BPNFS11	2.733
CEI3	and	BPNFS23	2.980
STPS5	and	POQ1	2.559
STPS5	and	POQ2	2,814
CEI1	and	POQ2	2.576
AGQR11	and	POQ2	2.549
CEI5	and	STPS1	2.778
CEI1	and	STPS1	3.816
AGQR12	and	STPS5	3.017
CE1	and	BPNFS19	3.017
AGQR5	and	CEI1	2.686
AGQR11	and	CEI1	5.593

AGQR10	and	CEI1	3.037

Table 65 presents the standardized residuals of the current study. Due to space considerations, only residuals greater than ± 2.5 were included. Out of 703 residuals only 27 residuals (eight negative and nineteen positive) were out of the suggested threshold, which is slightly above 3.84% of the total residuals. Three of the six negative residuals were associated with Scarcity dimension of the model which also has one of the lowest factor loadings. Another three highest negative residuals were related to DMC construct. When positive residuals were investigated, it could be seen that six of them were again associated with Scarcity dimension. In addition, there was only one standard residual above 4.0 which is between AGQR11 and CEI1. Nevertheless, as Hair et al. (2013) point out and based on the good fitness and adequate construct validity, it was decided to retain these poorly performing items to satisfy the statistical requirements.

So far it could be seen from the above analyses, the proposed model by Chou (2015) fits the data and it was found out that factors had significant correlations. The last model diagnostic to be checked is the modification indices which are calculated for each fixed parameter in the model. Hair et al. (2013) point out that among the modification indices obtained from the analyses, the two sets of indices are modification indices for factor loadings and for the error terms (p. 636).

Due to space considerations, full table of modification indices (MIs) was not given. Instead only the largest indices were examined. When the covariances of the error terms were investigated, it was seen that the largest MI was 23.462 for the covariance of the error terms of CEI7 and CEI9. Then the second largest MI was 20.826 for the covariance of the error terms AGQR6 and AGQR3. However, Meydan and Şeşen (2015) argue that when the goodness-of-fitness indicates a satisfying fitness, and there may not be a need to apply any modification based on MIs (p. 82). Similarly, Hair et al. (2013) remark that any model respecifications applied by correlating error terms should be avoided. Thus, no respecification was applied for the model based on the MIs.

In addition to MIs between error terms, MIs for factor loadings which were provided in Table 66 were also examined. As can be seen, all constructs have MIs for each item except for the items aimed at measuring them.

The largest MIs were 23.74 between CEI1 and CD8a, 11.866 between STPS5 and CD8b, 11.283 between STPS5 and CD4, and 11.144 between CEI1 and CD1. The high MIs related to these items might be resulting from low factor loadings.

Table 66

Modification Indices for Factor Loadings for CFA Model of Octalysis

	Items	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8a	CD8b
CD1	DMC1	-	.087	.126	.149	.412	.08	Oca.14	.258	4.064
	DMC2	-	.303	.644	2.213	.705	.164	2.166	1.311	.93
	DMC5	-	5.684	4.179	.829	.006	.175	.23	2.689	2.867
	DMC6	-	3.488	4.065	1.268	3.029	.03	.073	3.125	0
	DMC7	-	2.171	1.542	.146	.09	0	.495	.571	1.197
	DMC12	-	.825	4.972	5.679	.544	.42	.108	.143	.783
CD2	BPNSF5	.11	-	2.685	6.676	0	.382	.059	3.307	1.963
	BPNSF11	1.751	-	7.037	4.92	1.598	1.917	.002	.639	9.643
	BPNSF17	.654	-	1.575	6.658	.06	.288	2.539	.373	2.685
	BPNSF23	2.779	-	4.469	4.644	.832	1.632	1.837	3.094	6.042
CD3	BPNSF1	3.989	.555	-	1.546	7.725	5.471	.612	2.312	1.948
	BPNSF7	3.171	1.216	-	3.131	6.969	1.598	.239	.771	.283
	BPNSF13	1.838	.919	-	.55	.472	1.976	2.161	1.252	.027
	BPNSF19	4.196	.421	-	3.528	.095	1.558	4.11	103	.144
CD4	POQ1	.863	.565	1.491	-	2.889	.017	.857	2.498	.206
	POQ2	.024	1.821	3.913	-	.743	1.345	7.108	2.85	2.382
	POQ3	.054	.002	.176	-	.594	2.035	2.018	.048	5.439
	POQ5	.754	.105	.339	-	.647	.047	.074	.359	.678
CD5	BPNSF3	6.847	4.984	2.061	.747	-	.393	4.44	1.682	.442
	BPNSF9	.454	1.609	.206	.468	-	.009	6.821	1.042	.077
	BPNSF15	.2	.818	.358	.891	-	.139	.08	.059	1.675
	BPNSF21	1.774	.101	.013	.009	-	1.676	.612	.505	2.418
CD6	STPS1	1.709	4.208	5.572	.064	2.053	-	7.318	.052	6.396
	STPS3	.948	.38	.017	6.862	.804	-	1.196	.096	1.466
	STPS4	1	4.171	10.609	2.176	1.205	-	2.124	.161	.248
	STPS5	.102	1.715	.096	11.283	2.572	-	8.135	.039	11.866
CD7	CEI1	11.144	.853	5.873	5.044	4.579	8.478	-	23.74	8.688
	CEI3	3.537	3.441	.22	7.659	1.259	5.635	-	8.215	1.555
	CEI5	1.729	.455	2.481	5.821	8.19	7.028	-	4.974	.579
	CEI7	.422	6.985	6.099	3.033	10.305	5.603	-	7.658	1.905
	CEI9	2.598	.839	1.074	.129	1.105	.884	-	1.273	.51
CD8a	AGQR5	.811	.76	.032	.21	.266	.848	2.275	-	.54
	AGQR9	5.647	4.869	3.876	1.698	1.167	1.492	4.882	-	.959
	AGQR11	8.071	2.939	8.928	6.874	5.414	.899	3.577	-	.584
CD8b	AGQR6	.02	.059	.077	.137	.113	.138	.238	4.606	-
	AGQR10	1.219	.186	.085	3.025	.743	.309	.288	.003	-

AGQR12	1.462	.04	.001	1.859	1.326	.041	.006	3.659	-

* CD1: Epic Meaning and Calling; CD2: Development and Accomplishment; CD3: Empowerment of Creativity and Feedback; CD4: Ownership and Possession; CD5: Social Influence and Relatedness; CD6: Scarcity and Impatience; CD7: Unpredictability and Curiosity; CD8a: Loss and Avoidance (Mastery Avoidance), CD8b: Loss and Avoidance (Performance Avoidance)

When the rest of the Table 66 was examined, it could be seen that apart from these largest MIs, there were 62 more MIs of 4.0 or higher, most of which were slightly higher than the suggested threshold. Based on these findings, it was decided that there was no need for model respecification because all the fitness indices indicated good fit to the data and the model was sound in terms of its theoretical basis.

So far, with three different data sets, EFA analyses and CFA procedures were run and the theoretical model supposed by Chou (2015) was validated. These consisted of the first part of the process of Structural Equation Modelling (SEM), which is also called as measurement model. The second part of SEM requires the specification of structural model which includes a visual diagram of structural equations to show the relationships in the theoretical model. In the measurement model the focus is on the relationships between latent constructs. On the other hand, in SEM the focus shifts to the magnitude of these relationships (Hair et al., 2013). In SEM part, to accept the proposed model, its overall and relative fitness are tested and then parameter estimates are examined. If there is a good fit and hypothesized paths are significant, it is possible to state that there is enough support for the model (Hair et al., 2013; p. 642).

For SEM, first the findings obtained from the CFA procedure need to be transformed into a structural model. What is aimed with structural model transformation is to change the correlational relations in CFA to structural model relationships along with some other changes such as notation changes. To conduct theoretical transformations, two-headed arrows indicating correlational relationships between the construct should be replaced with single-headed arrows indicating hypothetical relationships between those constructs. This results in two different types of construct: endogenous which are the ones act as outcomes and at which single-headed arrows point. Since endogenous constructs are not fully explained they are added error terms.

In the process of modelling, it is possible to apply different approaches. For example, a path analysis can be done. To perform a path analysis, fist the measurement model need to be validated and then, composite variables are formed. In this way latent variables can be turned into observed ones. By forming composite variables for each construct, it is possible to run the path analysis to examine the relationships between constructs (Collier, 2020; p.128). In a path mode, only squares are used to represent the relationships. However, as Collier (2020) and Hair et al. (2013) state, using path analysis brings about some problems. It does not represent measurement error, which makes it difficult to explain the variance in a model. Thus Collier (2020) suggests using full structural models, which allows accounting for the measurement error of indicators while examining the relationship between constructs. In the same vein, Hair et al. (2013) also suggest using full structural models by using the factor patterns obtained in CFA. In this way, it becomes possible to include error terms to be estimated. And it is possible to make a comparison between the fitness of CFA model and structural model. Therefore, in the scope of this study, full structural model was tested.

Stage 5: Specifying the structural model. The specification of full structural model starts with the definition of the structural theory: Octalysis. The Octalysis theory (Chou, 2015) hypnotizes that CD2: Development and Accomplishment; CD3: Empowerment of Creativity and Feedback; CD4: Ownership and Possession; CD5: Social Influence and Relatedness; CD6: Scarcity and Impatience; CD7: Unpredictability and Curiosity; CD8: Loss and Avoidance are related to CD1: Epic Meaning and Calling. Accordingly, activities which aim at improving creativity and making individuals more autonomous; which allow learners to get into interaction with their friends and foster group activities; which develop a sense of ownership during in learning processes and environment; and which let individuals feel they are improving their language skills make it possible to achieve long term motivation. In addition, utilizing scarce items or limitation of sources, and getting individuals strive to obtain goals just not to lose something; or integrating new and unexpected things into the language learning process to make individuals get into action because of their curiosity can help to maintain language learning motivation in the long run. In this framework, this study aimed at testing the following hypotheses:

English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with:

1. their feeling of accomplishment, skills development and overcoming challenges
2. their involvement in creative processes
3. their feeling of the ownership regarding their own learning processes
4. their feeling of being related to a social group in their learning environments
5. their desire to obtain scarce things and their perception of closing deadline
6. their being curious about the things in their learning environments which get their attention and their encounter with unpredictable things
7. their endeavors not to lose something and to avoid failure

The visual diagram of the theory is shown in Figure 14.

As can be seen in the model, all the structures except for CD1: Epic Meaning and Calling were exogenous and they were utilized to predict CD1. Mastery Avoidance and Performance Avoidance which were presented as CD8a and CD8b in the CFA analysis were grouped under the heading Loss and Avoidance. In this respect a second order construct was created. The constructs, CD3: Empowerment of Creativity and Feedback; CD2: Development and Accomplishment; CD5: Social Influence and Relatedness; CD4: Ownership and Possession; CD7: Unpredictability and Curiosity: CD6: Scarcity and Impatience; and CD8: Loss and Avoidance were all considered to be correlated with each other so two-headed arrows were used to covariate them. On the other hand as Chou (2015) indicates all of the above mentioned constructs are used to improve learners' long-term bigger motivation, which is Epic meaning and Calling. So, this construct was shown as endogenous one which was pointed by single-headed arrows from the other constructs. The structural model presented in Figure 14 can be estimated and assessed.

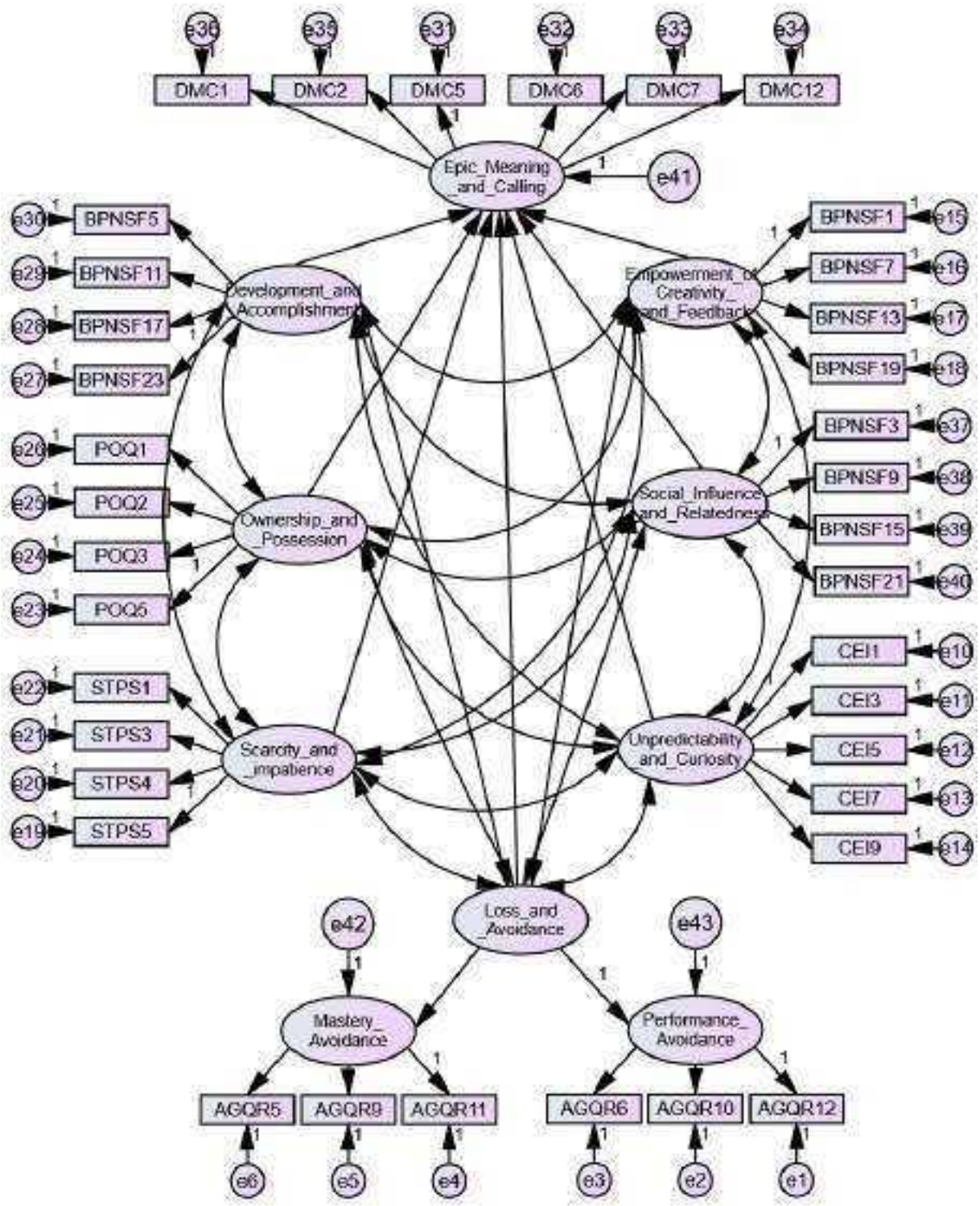


Figure 14. Structural model of Octalysis

Stage 6: Assessing the Structural model Validity. In Table 67, overall fitness statistics for the structural model for Octalysis are given together with the results of CFA to allow the comparison of the results of both analyses.

Table 67

Comparison of Goodness-of-Fit Measures Between Structural Model and CFA of Octalysis

GoF Index	CFA Model	Octalysis Model
Absolute Fit Measures		
Chi-square (CMIN)	1243.744	1272.190
Degrees of freedom (DF)	593	599
Probability (p)	.000	.000
CMIN/DF	2.097	2.124
Goodness-of-fit index (GFI)	.945	.944
Root mean square error of approximation (RMSEA)	.031	.031
90 percent confidence interval for RMSEA	.028; .033	.029; .034
Root mean square residual (RMR)	.039	.042
Squared root mean square residual (SRMR)	.0363	.0382
Incremental Fit Measures		
Normed fit index (NFI)	.908	.906
Comparative fit index (CFI)	.949	.948
Relative fit index (RFI)	.896	.895
Parsimony Measures		
Adjusted goodness-of-fit index (AGFI)	.935	.934
Parsimony normed fit index (PNFI)	.808	.815

As it is shown in Table 67, the x^2 was 1272.190 with 599 degrees of freedom at p = .000 significance level. The normed chi-square was 2.124. When other fit indices were controlled, it could be seen that CFI was .948 with a RMSEA of .031 (90% confidence interval of .029 to .034). All of these measures were within the suggested limits and indicated an overall good fit. When these results were compared to the results of CFA, little changes were observed, which could be considered insignificant.

In the next step, standardized estimates of the loadings and path coefficients were investigated (Hair et. al., 2013; p. 656). By using the data provided in Table 67, it was possible to make a comparison of CFA results and SEM results in terms of loading estimates of the items. It was observed that the loading estimates did not substantially change in SEM model. There were only minor changes in the items of Scarcity construct and Mastery avoidance. The loading of STPS 4 did not change. On the other hand, the loading of STPS1 increased from .55 to .57, which was the highest difference. STPS3 which was .54 previously dropped to .53, and STPS5 which was .54 dropped to .53. regarding the changes in Mastery Avoidance dimension, the loading estimate of AGQ9 did not change whereas AGQ 5 whose loading estimate was .58 dropped to .57 and AGQ11 whose estimate was .53 dropped to .52. Any other change in the loading estimates of the other items was

not observed. This finding indicated that parameter stability was achieved. In addition, as Hair et al. (2013) point out, stability in parameters could provide evidence for validity of the measurement model (p.656). Table 68 provides findings of the comparison of standardized factor loadings and construct reliabilities for structural and CFA models of Octalysis.

Table 68

Comparison of Standardized Factor Loadings and Construct Reliabilities for Structural and CFA Models of Octalysis

	Variable Label*																	
Variable Name	CD1		CD2		CD3		CD4		CD5		CD6		CD7		CD8a		CD8b	
	CFA	SEM	CFA	SEM	CFA	SEM	CFA	SEM	CFA	SEM	CFA	SEM	CFA	SEM	CFA	SEM	CFA	SEM
DMC1	.66	.66																
DMC2	.54	.54																
DMC5	.60	.60																
DMC6	.68	.68																
DMC7	.41	.41																
DMC12	.60	.60																
BPNSF5			.66	.66														
BPNSF11			.66	.66														
BPNSF17			.69	.69														
BPNSF23			.71	.71														
BPNSF1					.54	.54												
BPNSF7					.66	.66												
BPNSF13					.73	.73												
BPNSF19					.61	.61												
POQ1							.92	.92										
POQ2							.80	.81										
POQ3							.72	.72										
POQ5							.83	.83										
BPNSF3									.53	.53								
BPNSF9									.50	.50								
BPNSF15									.73	.73								
BPNSF21									.52	.52								
STPS1											.55	.57						
STPS3											.54	.55						
STPS4											.39	.39						
STPS5											.54	.53						
CEI1													.53	.53				
CEI3													.63	.63				
CEI5													.71	.71				

CEI7													.58	.58				
CEI9													.68	.68				
AGQR5															.58	.57		
AGQR9															.77	.77		
AGQR11															.53	.52		
AGQR6																	.80	.80
AGQR10																	.83	.83
AGQR12																	.84	.84
Composite Reliability	.77	.76	.78	.78	.73	.73	.90	.89	.66	.66	.58	.58	.76	.76	.66	.66	.86	.86

* CD1: Epic Meaning and Calling; CD2: Development and Accomplishment; CD3: Empowerment of Creativity and Feedback; CD4: Ownership and Possession; CD5: Social Influence and Relatedness; CD6: Scarcity and Impatience; CD7: Unpredictability and Curiosity; CD8a: Loss and Avoidance (Mastery Avoidance), CD8b: Loss and Avoidance (Performance Avoidance)

As can be seen from Table 68, the consistency between the loading estimates of CFA and SEM analyses resulted in consistency in composite reliabilities. Mostly there were no changes in terms of the composite reliabilities of the constructs. There was .01 decrease in CD1: Epic Meaning and Calling and CD5: Ownership and Possession, which could be ignored.

To determine the validity of the model, Hair et al. (2013) also state that significance of the path estimates should be checked. Table 69 presents findings in terms of structural parameter estimates for Octalysis model.

As can be seen from Table 68, all the estimates were in the expected direction and apart from the estimate between CD3 and CD1, all other path estimates were significant at .05 level. Nevertheless, since six estimates out of seven were significant enough in line with the hypotheses, it would not be wrong to assume that the theoretical model was supported.

Table 69

Structural Parameter Estimates for Octalysis Model

	Structural Relationship			Unstandardized Estimate	S.E.	T Value	Standardized Estimate	Hypothesis Supported?
H1	CD2	→	CD1	.161	.069	2.33	.168	Supported
H2	CD3	→	CD1	.024	.071	.34	.023	Not Supported
H3	CD4	→	CD1	.093	.025	3.773	.149	Supported
H4	CD5	→	CD1	.106	.053	2.003	.094	Supported
H5	CD6	→	CD1	.156	.048	3.27	.183	Supported

H6	CD7	→	CD1	.146	.073	1.983	.113	Supported
H7	CD8	→	CD1	.102	.049	2.089	.091	Supported

* CD1: Epic Meaning and Calling; CD2: Development and Accomplishment; CD3: Empowerment of Creativity and Feedback; CD4: Ownership and Possession; CD5: Social Influence and Relatedness; CD6: Scarcity and Impatience; CD7: Unpredictability and Curiosity; CD8a: Loss and Avoidance (Mastery Avoidance), CD8b: Loss and Avoidance (Performance Avoidance)

Although goodness-of-fitness indices and path estimates yielded a satisfying model that supports the theory, model diagnostics were examined to determine whether there would be any need to respecify the model. For this purpose, standardized residuals and modification indices of the structural model were investigated. Table 70 provides information regarding the standardized residuals higher than ±2.5.

Table 70

Model Diagnostics for Octalysis Model

Standardized residuals (residuals greater than ± 2.5)			
Negative standardized residuals			
CEI9	and	BPNSF9	-2.611
CEI7	and	BPNSF9	-3.58
AGQ12	and	DMC1	-2.549
CEI9	and	DMC2	-2.716
DMC12	and	DMC12	-2.509
DMC7	and	DMC12	-2.583
AGQR6	and	BPNFS23	-2.886
AGQR10	and	BPNFS23	-2.872
CEI7	and	STPS5	-3.195
AGQR6	and	CEI3	-2.61
Positive standardized residuals			
STPS3	and	BPNFS21	2.553
CEI1	and	BPNFS21	2.787
CEI1	and	DMC1	3.03
DMC2	and	DMC7	2.78
BPNFS19	and	DMC12	2.802
BPNFS1	and	DMC12	3.054
BPNFS19	and	BPNFS11	2.733
CEI3	and	BPNFS23	2.971
STPS5	and	POQ1	2.95
AGQR11	and	POQ2	2.536
STPS5	and	POQ2	2,772
CEI1	and	STPS1	3.675
CEI5	and	STPS1	2.595
AGQR6	and	STPS5	3.63
AGQR10	and	STPS5	3.672
AGQR12	and	STPS5	4.619
AGQR11	and	BPNFS1	2.684
AGQR5	and	CEI1	2.967
AGQR11	and	CEI1	5.867
AGQR9	and	CEI1	2.655
AGQR6	and	AGQR5	2.514

When Table 70 is examined, it could be seen that it was not possible to diagnose a pattern between specific constructs that signaled a need for model respecification. Moreover, when model modification indices were investigated, it was found that there were not high enough modification indices in terms of the relationships between the structural constructs. That is, it would not have been possible to improve the model substantially, if the suggested modifications had been applied. Therefore, it was decided to retain the model as suggested.

All in all, the structural model that supported the hypotheses based on the Octalysis model had a good but not perfect overall fit.

Phase 2

The primary objective of the current study was statistically testing the validity of Octalysis framework to improve the long-term motivation of language learning students theorized by Dörnyei (2016) as Directed Motivational Currents. The secondary objective, on the other hand, was to design a gamification application based on the Octalysis model that could be utilized in real language learning environments. In this respect, in the second phase of the study quasi-experimental convergent parallel-mixed method was adopted. This phase consisted of two stages. At the first stage, it was aimed to collect quantitative data and then at the second stage in which qualitative data were collected the focus was on the application of Octalysis based gamification design.

Participants. For the second phase of the study, two classes which had been attending one-year compulsory language preparatory program for six months in School of Foreign Languages (SFL) in Muğla Sıtkı Koçman University (MSKU) were randomly selected by the researcher. All the students were informed about the aim and the content of the study and they were asked to give their written and oral consents. But before that, all the students in the two classes were informed that the study which they had been invited to participate was a voluntary and it was possible to leave the study any time they wanted throughout the study or demand their data collected during the study to be excluded after the study. Moreover, the students were assured that the data were only collected for research purposes declared in the written format and would not be used for other purposes and/or shared with third parties. Finally, students were given contact numbers and email addresses in case

they might have had questions and/or desired to be informed about the results of the study. Eventually, 20 students from the classroom one which was assigned to experimental group, and 15 students from the classroom two which was assigned to the control group gave their consents both in written and oral format and accepted to participate into the study.

The above-mentioned procedure was carried out in the first stage of the second phase with all the students who accepted to take part in the study. Besides, the students had also been invited to part in the interviews which were to be held before, through and after the gamification intervention. Five students from the experimental and five students from the control group accepted the invitation. These ten students were also informed that participating into the interviews was totally voluntarily and they would leave the study any time they wanted and/or asked for exclusion of their data that was supposed to be collected in the interviews. In addition, the students were informed that the interviews were to be audio-recorded and the recordings wouldn't be used apart from research purposes and/or shared with third parties, and would be deleted afterwards the study. Finally, they were asked to give their written and oral consents.

At the time of the study, in MSKU all the students enrolled in 100% English medium programs had to either attend one-year compulsory language preparation programs in SFL or prove that they had required level of language demanded by the related program. Before starting the academic year, each student had to take a diagnostic test that was administered by SFL, then, the students were placed into three groups in accordance with their results. The first group of students who got the lowest marks would be placed into a program in which they took 24 hours of language education for 30 weeks. The second group of students who got relatively higher scores but unable to pass B1 level were placed into the program in which they got 20 hours of language education for 30 weeks. And finally, the third group of students mostly consisted of the ones who had achieved B1 level in their previous education placed into the third program in which they got 18 hours of language education for 30 weeks.

In addition to students, data were also collected from the teachers of the both classes. The teachers were also informed about the process and their rights as the participants. They were told they were free to leave the study any time they wanted

and/or demand their data to be deleted after the study. In addition, they were assured that their data would be audio-recorded and would not be used except for the research purposes and/or shared with third parties. Then, the teachers were also asked to give their consents in written form by signing the consent paper and orally.

The students both in the experimental and control groups were among the students who had been enrolled into 100% English medium programs in MSKU and who were placed in second program in SFL according to their diagnostic test results that had been carried out at the beginning of the 2018-2019 Academic Year, which meant that the proficiency levels of the students were more or less the same. In addition, all the students were attending classes for the last six months.

Data collection process

The quantitative data at the first stage were collected via the survey obtained in the first phase of the study. The survey was administered in both control and experimental groups with the participation of 35 students. The findings at this stage were used to define Octalysis strategy dashboard (Chou, 2015; p. 463) and in turn Octalysis for each classroom. Researcher also collected interview data from semi-structured audio-recorded weekly interviews. Thus, it would be possible to get participants' opinions about the intervention. In addition, the teachers of both classes were asked to keep logs and keep records of their classrooms and their students' reactions in terms of gamification application. Moreover, pre and post audio-recorded interviews were also carried out with the teachers, which constitute another data source.

Table 71

Data Collection Steps in 1st and 2nd Stages of the Second Phase

1st Stage	2nd Stage
• Survey application	• 7-week Intervention
• Pre-Intervention Interviews	• Weekly Interviews
• Drawing Octalysis Framework	• Post-Intervention Interviews
• Defining Octalysis Strategy Dashboard	• Teachers Logs

Data analyses (t-tests)

As the first step in the second phase of the study, the students in both classrooms were asked to complete Octalysis survey which had been validated in the first study. A total of 35 students (20 students from experimental group and 15 students from the control group) completed the survey. Students' responses were collected and for each dimension of the Octalysis model, the mean scores were calculated. Then to test whether the two groups, experimental and control, significantly differ from each other, independent sample t-tests, whose results were presented in Table 72, were run.

Table 72

Independent Samples t-Test Results

CD	Group	N	Mean	SD	t	df	Sig. (2-tailed)	Mean Difference
Epic Meaning and Calling	Experimental	20	3.84	1.07	.013	33	.990	.004
	Control	15	3.83	.75				
Development and Accomplishment	Experimental	20	6.03	1.16	-.116	33	.908	-.058
	Control	15	6.08	1.81				
Empowerment of Creativity and Feedback	Experimental	20	5.53	1.01	.471	19.74	.643	.258
	Control	15	5.27	1.93				
Ownership and Possession	Experimental	20	3.70	.90	-.054	33	.958	-.017
	Control	15	3.72	.93				
Social Influence and Relatedness	Experimental	20	6.58	.90	.757	20.98	.458	.342
	Control	15	6.23	1.56				
Scarcity and Impatience	Experimental	20	6.36	.88	.127	33	.900	.045
	Control	15	6.32	1.26				
Unpredictability and Curiosity	Experimental	20	5.51	1.43	-.475	33	.638	-.250
	Control	15	5.76	1.68				
Loss and Avoidance	Experimental	20	5.22	1.51	-.416	33	.680	-.228
	Control	15	5.44	1.73				

As can be seen in Table 72, Findings indicated that there was not any significant difference between the experimental (M = 3.84, SD = 1.07) and control (M = 3.83, SD = .75) in terms of the effect of CD 1: Epic Meaning and Calling, t (33) = .013, p = .990, which meant the students had the same level of goal pursuit in both groups. Secondly, there was not any significant difference between experimental group (M = 6.03, SD = 1.16) and Control group (M = 6.08, SD = 1.81) in terms of the effect of CD 2: Development and Accomplishment, t (19.74) = -.116, p = .908. This meant that students from both groups had similar tendencies regarding their feeling of accomplishment, skill development and overcoming challenges.

Next independent t-test was carried out to examine the effect of CD 3: Empowerment of Creativity and Feedback, which included involvement in creative processes and aiming at becoming autonomous. The findings indicated that there was no significant difference between experimental (M = 5.53, SD = 1.01) and control (M = 5.27, SD = 1.93) groups in terms of the effect of CD 3, t (33) = .471, p = .643. In addition, in another independent sample t-test, it was found out that there was not any significant difference between experimental (M = 3.70, SD = .90) and control (M = 3.72, SD = .93) regarding the effect of CD 4: Ownership and Possession, t (33) = -.054, p = .958. this meant that students in both groups felt ownership at the same level.

The findings also confirmed that there was not any difference between the experimental (M = 6.58, SD = .90) and control (M = 6.23, SD = 1.56) groups in terms of the effect of CD 5: Social Influence and Relatedness, t (20.98) = .757, p = .458. Moreover, it was found out that CD 6: Scarcity and Impatience had similar effects on students in experimental (M = 6.36, SD = .88) and control (M = 6.32, SD = 1.26) groups, t (33) = .127, p = .900. These findings indicated that students interacted with their friends and engaged into the activities at the same level and they also perceived the roles of scarce and limited things in their language learning environments at the same level.

In the final two independent sample t-tests, it was found out that there was no difference between the experimental (M = 5.51, SD = 1.43) and control (M = 5.76, SD = 1.68) groups in terms of the effect of CD 7: Unpredictability and Curiosity, t (33) = -.475, p = .638. In a similar vein, there was no difference between experimental (M = 5.22, SD = 1.51) and control (M = 5.44, SD = 1.73) in terms of the effect of CD 8: Loss and Avoidance. Based on these findings it could be assumed that the perception of new and interesting items that made students curious were at the same level and students from both groups had avoidance motivation at the same strength.

The other data sources at this stage were the audio-recorded interviews with students and teachers. Semi-structured interviews had been carried out with ten volunteered students (5 from experimental group and 5 from control group) and three teachers, and they had been asked to evaluate the language preparatory

programs that they had been attending for the last six months in terms of the eight core drives of Octalysis.

Pre-Intervention Teacher Interviews, General Evaluation

In scope of the study, to make a holistic analysis and general evaluation of the language learning environment teachers giving lectures in experimental and control groups were asked to state their opinions regarding the language learning and teaching processes at that time. First teachers were informed about what the Octalysis was and the core drives that made up Octalysis. In addition, the underlying theories such as Flow Theory, Approach and Avoidance Theory, Intrinsic and Extrinsic Motivation and Directed Motivational Currents had been briefly told. Then, an explanation about how Octalysis and its core drives can be applied in language learning environments was made.

CD 1: Epic Meaning and Calling. Teachers of both classrooms were first evaluated their classrooms in terms of their and their students' goals, which was CD 1. Instead of evaluating each activity and/or material the teachers used in their classrooms in detail, they were asked to make in a rather holistic way. According to the findings from their interviews, it could be said that it was not possible to mention the existence of totally meaningful and bigger goals for every student in both classrooms. There are a few students who have these kinds of goals, though. Teachers stated that many of their students not try to achieve big goals but try not to fail, which was a typical example of approach and avoidance motivation. A teacher of experimental group said that:

> Extract 1. In terms of the having meaningful goals, when I evaluate all the things going around in my classroom, it is for sure that many of my students are studying not to fail in the exam. That is, they obviously try to avoid failure in language preparation program (T1)
>
> Extract 2. I know that many students in my classroom do not have any goals. To pass the exam is just enough for them. They frequently ask me what the minimum grade is to be accepted as successful (T3).

However, teachers added that there were some students who really pursued bigger goals. There were students who really wanted to improve their foreign language and had long-term dreams. Teachers indicated that those students were aware of the language learning processes. A teacher from experimental group stated that:

> Extract 3. There is a student in my classroom. That student is very interesting because when I asked him the reason for learning a language, he said that he wanted to be world citizen (T1)

It was possible to find students with bigger goals in control group, too. A teacher from that group pointed out that students especially who had been attending engineering and medicine faculties were different from the students of other faculties and had big goals.

> Extract 4. Many of the students from these faculties know that they will have to use the target language. They want to have masters or a PhD degree abroad and they are aware of the fact that they should use the target language effectively (T3).

CD 2: Development and Accomplishment. For the second core drive CD2: Development and Accomplishment, the teachers were informed that when students feel that they are developing and improving their language proficiency levels, it is highly likely that they will be more motivated. In many gamification applications, researchers include the game elements such as points, badges and leaderboards so that students or participants can quantify their development and go on their studies. In addition, maybe one of the most frequently used strategies is to group students into proficiency levels in language learning classrooms. When students completed a level, they jumped to the next one and in this way, they could observe their development. In this respect, the teachers stated that their students could feel they were achieving something and followed their developments. And once they felt that they were achieving, they started to struggle for bigger aims. In this respect, teachers added that one of the most influential tools was formative assessments which were administered throughout the academic year. A teacher from the experimental group pointed out that:

> Extract 5. Most of the things that we do such as exams and quizzes, I think, make my students feel that they feel they are developing and improving their language, especially hardworking ones (T2).

A similar comment was made by a teacher of control group:

> Extract 6. They like following their grades online. I think they just love feeling of achieving something. But some of them really follow their development. They keep their records (T3).

Teachers did not only focus on the grades in this sense, they also indicated that some of their students counted the vocabulary items they had memorized by keeping vocabulary books. In addition, a teacher from control group said that he/she

always kept the first pieces of students' writings. Then, when he/she felt that students were having trouble, he/she would show it so that the student could see the difference. As a result, in terms of CD2, teachers of both classrooms stated their classrooms include elements related to CD2 and they gave higher point for CD2.

CD 3: Empowerment of Creativity and Feedback. CD3: Empowerment of Creativity and Feedback was related to the creative power of language learning activities that had been applied until that time. Teachers were told that when CD3 for a classroom was high enough, students started to enjoy just for the sake of learning something and they started to become more autonomous learners who could take their own responsibilities. A teacher from experimental group indicated that there were students who were really creative.

> Extract 7. I have students who are really creative, they mostly do what they need to do and bring me extra activities they do outside the classroom (T2).

Another teacher from control group said he/she try to promote creative activities and make students become autonomous ones, which was actually at the core of this drive.

> Extract 8. I especially try to pay attention to stimulate students' creative sides to be triggered by the activities I do. I don't like simple ones. And I give extra credit for their extra-curricular activities (T4).

A different comment came from another teacher of control group that proved the existence of creative processes in control group.

> Extract 9. From time to time, I stop the lesson and I try to show my students the enjoyable part of language learning. And I want them share their favorite experiences in this sense (T3).

Consequently, teachers of the both classrooms give similar points for CD3 for their language learning environments and activities they used.

CD 4: Ownership and Possession. On the Octalysis framework, Ownership and Possession was the fourth core drive. Chou (2015) argues that once individuals begin to feel the sense of ownership in terms of their own learning experiences, they become more motivated to improve it. In addition, they try to protect or get more of it. That's why teachers want their students internalize the subject they are working on and ask their students integrate their real-life experiences into learning environments. It is possible to see the effect of using CD4 even in many language learning environments. For example, many teachers try to set up the rules of their

classrooms together with their students and they want them to be active members. What they try to do is to make students feel the ownership of their classrooms, which would make them try to protect and improve. It is possible for a teacher see that students prefer to collect many things such as their worksheets, badges or small stickers. The reason behind this behavior is not just to collect but to feel they that own all those items, which in turn lead them to trying to protect them.

Teachers from experimental group indicated that they made use of stickers and badges, which were provided by the publisher of the main course book in his/her group.

> Extract 10. I use those badges behind the book. It's funny that I photocopy them and cut with scissors by myself. I mean they are not valuable. But students love them. You could see how some students tried to get them (T2).

However, the teacher in control group thought that CD4 should be low in his/her classroom because some students even do not want to be a part of school of foreign languages, indicating that those who did not want to be at MSKU disturbed the atmosphere in terms of CD4.

> Extract 11. Since this is language preparation classroom, feeling ownership is a little bit problematic and there are students who stuck in the middle (T4).

As a result, it could be said that although the teachers from the experimental and control groups slightly graded differently their classrooms in terms of CD4, it was still possible to mention a balance.

CD 5: Social Influence and Relatedness. The next core drive in Octalysis framework was CD5: Social Influence and Relatedness. As its name suggests, this core drive is about the relationships between students in a language learning environment. For a long time, it has been known that social dimensions of language learning are very important, and it is known that students perform better when they act in a group and spend time on activities together with their friends. When the teachers of experimental and control groups asked to evaluate their classrooms in this respect, differences were observed. A teacher from experimental group found his/her classroom environment weak in terms of students' social relatedness.

> Extract 12. From time to time I integrate group activities. Sometimes it works sometimes not. I don't know why (T2).

On the other hand, a teacher of control group stated that although from time to time negative consequences of social dimension could be seen, generally group activities were preferred.

> Extract 13. It is for sure that students are affected by their peers. I sometimes try to soothe peer pressure but sometimes there are students who are affected negatively. But I can say that I make use of the group activities as well. I believe that they like it (T3).

Another teacher from control group commented in a parallel way. And said that

> Extract14. Social relatedness is not so bad. Students have their own groups and can work together (T4).

As to the teachers' evaluations differed most on CD5. It was found out that in experimental group teachers thought that social elements and activities were not so effective whereas teachers from control group believed that their students perform better with their peers.

CD 6: Scarcity and Impatience. One of the black hat drives that should be used very carefully is CD6: Scarcity and Impatience. This drive indicates scarce things which are difficult to obtain or not abundant cause individuals to get into action. Because when things become scarce, people start to perceive that they are more valuable. It is possible to see many applications of CD6 in language learning classrooms. For example many teachers offer small prizes for the students who finish first, or they just appreciate the students who acted faster. This means that not everybody could get the prize or appreciation from the teacher and they get into action immediately to get them. Another example is that many teachers limit the activity time, which means they apply CD6. After this short notice about the role of CD6, teachers evaluated their classrooms. Again, it was seen that there was not such a big difference between the two groups. A teacher from experimental group said that:

> Extract 15. I use scarcity, I usually offer small presents like chocolate as prize for the first and second students to finish the task. Can you believe that even they are university students they like it much (T1)?

On the other hand, a teacher from control group stated that he/she hardly ever use CD6 in classroom.

> Extract 16. I rarely use scarcity in my classroom due to the proficiency level of my classroom. Also, my students do not like competition. So, I don't like to use scarcity (T3).

However, another teacher from the control group stated on the contrary.

> Extract 17. When I think about scarcity, I can say that I often limit the time allocated for the activities. If it is a critical activity I offer extra points for students who hand in their assignment before due time (T4).

Depending on the above-mentioned extracts and teachers' opinions, it could be said that the two groups had close points although in experimental group it was more likely to see actions of CD6.

One of the other two core drives was CD7: Unpredictability and Curiosity. The role of chance is what determines the attractiveness of the activities. It is possible to find myriad activities that utilize curiosity element in language learning processes. For example, in all gap-filling activities it is aimed to find out something uncertain. Or activities including hidden objects and quests could be given as other examples. There was a similar evaluation from the teachers in this respect. A teacher from experimental group stated that:

> Extract 18. Most of the time I try to cover the grammar structures and I reserve most of the class time for grammar exercises. Thus, I can't say I try to integrate curiosity elements into my classes (T2)

A teacher from the control group had a similar point of view.

> Extract 19. Actually, the books we use include many activities that have some kind of tricks to increase curiosity but frankly speaking, if it requires extra, I don't pay attention (T4).

CD 7: Unpredictability and Curiosity. The findings about CD7 showed that it was graded lower when compared to other CDS, which meant that teachers from the both groups did not think that curiosity did not play a major role in their classrooms.

> Extract 20. Our work load does not let us integrate new and interesting things easily. I usually try to follow the pacing (T1).

As can be seen, a teacher from the experimental group acknowledged that there were not many applications of novel and interesting things which could help students get motivated. The reason for this, the teacher stated that the density of the program.

> Extract 21. I don't think that students will be willing to see new things because they hardly stay in the classroom and usually prefer to be passive listeners (T3).

Another teacher from the control group indicated that the students were so demotivated that it would not be possible to get their attention even if he/she tried.

CD 8: Loss and Avoidance. The final CD was CD8: Loss and Avoidance. Any unvolunteered actions trigger CD8. If students are forced to take part in activities involuntarily, then it becomes very difficult to improve their long-term motivation. Another problem with CD8 is studying not to lose or in other words to avoid losing or failing. This CD was by far the strongest CD that could be found in classrooms according to the teachers. A teacher from the experimental group stated that

> Extract 22. When the task at hand gets complicated and require productive skills, nearly half of the students either do not participate or just use basic and simple language. When I ask why, they say it is enough for them (T1).

Doing the minimum and not struggling to achieve more was a major problem in terms of motivation in experimental group. But the situation was not so different in control group. A statement made by a teacher in control group summarized the situation in a very good way.

> Extract 23. I can say that for many students not to fail in the exam is enough. They do not like forcing themselves for more. It is really irritating because on the long run I know that they will be unsuccessful (T3).

As could be understand for a many students studying not lose their status and just passing the tests at the minimum level was very normal and enough. However, avoiding so much meant that moving away from bigger and meaningful goals, which could be the reason why teachers gave very low marks for CD8.

At the end of the interviews, teachers were asked to identify the skills which they thought to be problematic and on which gamification applications might be applied. Many of the teachers stated that among the basic language skills, productive ones which are writing and speaking were the most problematic ones. They also added that it had become very difficult for them to motivate their students to write or speak in English.

> Extract 24. Definitely writing and speaking activities. I really found it difficult to motivate my students and enjoy these kinds of activities. They disliked writing and most of them did prefer to get into interaction in English (T1).

> Extract 25. When I give writing assignments, students usually either hand in incomplete tasks or tend to give others' works as if they are theirs (T4).

As can be understood from the comments made by the teachers from both classrooms, out of four skills, writing was a problem from the teachers' point of view, followed by speaking.

Finally, following the interviews, the grades which were given for each CD by the teachers and which indicated the extent to which related CD was thought to be dominant in classroom environment by the teachers were collected, and mean scores were calculated for each CD. It is possible to see the mean scores in Table 73.

Pre-Intervention Student Interviews, General Evaluation

Semi-structured interviews with students from both experimental and control groups were conducted in four steps. First, they were asked to state their thoughts about what they were thinking about the education given in the school of languages. Then, they got a short briefing about the composition of Octalysis and its eight core drives. After that, students evaluated the language education in terms of the eight core drives of Octalysis, and finally, they did the same evaluation for the writing activities and classes.

General Evaluation of the Program. In general, students had mainly two kinds of views about the education in school of foreign languages. One group of them were not satisfied because they thought that the education was not versatile enough. In addition, the students in this group criticized the education and stated that grammar was given much more importance than necessary. They also make similar comments for speaking and listening activities. Student 1 said that

> Extract 26. We only have a book and the most important thing is just to finish the unit at hand. Generally, the most important things are listening and speaking when you look at. But we do not do much listening, maybe once. As to speaking, it's up to teacher. If the teacher wishes, we do speaking practice (S1).

As can be seen, students were aware of the importance of speaking and listening. Moreover, they indicated that writing had similar problems as well. That is, it would not be wrong to state that productive skills were not among the issues on

which teachers focused enough. On the other hand, another group of students were satisfied with the language education.

> Extract 27. Generally, I am content with the education. However, it would be better if teachers did use more enjoyable activities or used jokes in the classroom. And I also would like them to use English all the time during classes (S5).

As can be understood from the comments made by Student 5, a group of students evidently stated that they were satisfied with the quality of language education. Yet, these students were also uncomfortable with the monotonous atmosphere of their classes. They mainly find it boring and lacking funny elements that would break that monotony. Moreover, like the students who stated that they were not satisfied with the language education, these group of students mentioned the shortage of time allocated for speaking practice. As in the case of Student 5, they disapproved teachers' use of mother tongue too much. This could also be interpreted that grammar education had significant place in classroom because it was very easy to teach grammar without using the target language.

After getting students opinions in general, a short briefing in which the aim and the use of Octalysis and the eight core drives of it were explained. Then, in the third step, students were asked to evaluate their language education in terms of eight CDs of Octalysis. One by one each CD was evaluated by the students in both groups. After that, they were asked to give a number between 1 and 8 (1 very weak – 8 very strong), which indicated the strength of their opinions. Then, their scores were summed up and a unique mean score for each CD was obtained. The mean scores were provided in Table 73.

CD 1: Epic Meaning and Calling. The mean scores for CD1 was the same for experimental and control groups which was 4.6. Largely, students were positive about having meaningful and big goals. For example, ST1 stated to have bigger goals and said that:

> Extract 28. ... I would need English in my life generally. And my goal is to achieve C1 level. I want that when I start talking to foreigners, I would like them to think that I am a native speaker (S1).

Student 1 had an ultimate goal of to be able to speak English as a native speaker, and was aware of the fact that English was not only a tool that was required for a successful school life but also would be an important of their future life. With

similar point of view, other students from the control and experimental groups indicated that English would be an essential part of their future lives. That meant that these students had long-term goals of learning a new language. Some students even stated they had bigger goals. During their education they wanted to be individuals who could speak English fluently and be able to communicate with foreigners and get into interaction abroad.

> Extract 29. ... English is a global language and wherever you go you can communicate in English. All human beings are connected with each other (S3).
>
> Extract 30. ... my main goal in not the language itself. It is to be able to get other information by using the language ... (S4).

As can be seen from the above extracts, many of the students were aware of the fact that knowing a foreign language would make them world citizens who could use language with communicative purposes. Even a student (Student 5) mentioned establishing a multinational software company and this student stated that English would be a key element to have a multinational status, which could be considered as a bigger goal than just being able to communicate in English. These findings indicated that students had meaningful and big goals. When students asked whether the language education and the activities used by the teachers contributed positively to their meaningful goals, they mainly agreed that school helped them achieve their goals.

> Extract 31. ... education here in this school does not make me struggle for big and meaningful goals. My goal is to learn English well. But the education isn't much related ... (S6).
>
> Extract 32. ... I don't think there are activities to create meaningful goals ... (S7).

So, although there were many students with meaningful goals who believed that the education given in MSKU School of Foreign Languages facilitated the possibility of realizing their goals, there were students (Student 6 and Student 7) who thought on the contrary. All-in-all, it could be stated that before starting the gamification intervention, students from the two classrooms had meaningful and big goals to some extent.

CD 2: Development and Accomplishment. Second core drive students commented on which was Development and Accomplishment. As in the CD1 there was a balanced evaluation between control and experimental groups, with a mean

score of 4.8 for both groups. Most of the students participating in interviews stated that when they felt they were developing, they would be motivated.

> Extract 33. It depends, ..., there are people who leave school because they do not see they are developing, but there are also people who start studying when they see they are developing. ... personally, if I couldn't improve myself, after some point I would give up ... (S1).

Students like Student 1 indicated that feeling the development and accomplishment was an important part of going on studying. Students also added that the activities having been carried out in school of foreign languages helped them feel they were improving their language skills. When they were asked how they would know they had been developing and improving their language skills, they gave various indicators. Among mostly stated indicators were turning a skill into a strategy, comparison with peers, being able to get into contact with foreigners, positive feedback they got and exam grades.

> Extract 34. for example, after learning some grammatical structure, when I realize that I can use it without thinking much, then I feel that I am improving my proficiency (S1).

Student 1, for example, thought that he/she could know that his/her language proficiency was developing when the skill, grammar in this case, turned into a strategy. That is, if somebody could use one skill and/or a structure without thinking, it would be an indicator of improvement. There were other students with different views as well. Student 7 from the control group or Student 2 from the experimental group stated that, exam grades were their markers of development.

> Extract 35. ... yesterday I go 81 from the exam. I know that it is not such a high mark. Yet, that was the highest mark that I've got from a quiz so far. and I felt my English is improving. And frankly speaking I got motivated (S2).

> Extract 36. The only thing that make me feel that my English is improving is the exams ... but ... I mean exam grades are not reliable. That is, there is nothing but exam scores, which is very problematic (S7).

As can be seen from the above extracts, exam grades were perceived as the indicators of development and accomplishment. However, for some students like Student 2 they were not enough. But on the other hand, there were some students who felt that they could track their language development thanks to the exams scores but who at the same time knew that scores were unreliable.

For some other students like Student 3, their peers' performances were very important to check their improvement.

> Extract 37. ... when I compare myself to my friends, I think I proceed in a linear way. And I absolutely think that my English is developing. And this makes me think that I can make it and thus I go on studying (S3).

Student 3 from experimental group stated that progress in relation to the group of students is a sign of development. This finding showed that students' performances needed to be evaluated in comparison with their peers, which indicated the coexistence of different CDs. Most of the time, it is possible to see various CDs in action together. In this case, a mixture of CD 2: Development and Accomplishment and CD 5: Social Influence and Relatedness could be seen.

There were some students who were not satisfied with the role of education in their perception of development. Student 4 from the experimental group was not totally sure that the language education given in school of foreign languages helped feel development.

> Extract 38. ... I can't say it is 100% but yes I can. I feel that my English is getting better but below my expectations. I mean I am not satisfied with my progress. I could develop more but ... conditions are not motivating ... (S4).

As can be understood from the extract given above, although Student 4 felt a little improvement in terms of language proficiency, he/she pointed out that there were other factors affecting the motivation level. Similarly, another student stated the presence of problems in this respect.

> Extract 39. ... this is the problem I am experiencing right now. I don't think that I made a good progress. This may partly due to my lack of studying enough, but partly because I cannot get feedback from the school that shows my progress. I feel that I am doing the same things all the time (S6).

Student 6 from the control group found the feedback mechanisms that had been in use in the classroom environment were not good enough and made him/her feel underdeveloped. This finding again signified how important to create a social dimension which included effective feedback mechanisms. It could also be understood from this student's interview that the monotonous structures which probably consisted of traditional way of teaching and lack of new and interesting things might have a harmful effect on students' perception of development which was an important drive behind sustainable long-term motivation.

CD 3: Empowerment of Creativity and Feedback. As in the case of the two previous CDs, there was not a considerable difference between the experimental and control groups. The mean score for experimental group was 4.6 whereas it was

4.4 for control group. In relation with this CD, students' evaluations focused on how a classroom could have a creative environment and how the activities used throughout language learning and teaching processes help students take their own learning responsibility.

> Extract 40. Teachers either totally focus on the book or the things they want, but not on us. Generally, we have monotonous classroom. Thus, there is not much creative things. This style doesn't help me use my creative powers (S1).

The findings from the above extract implies how important to create pleasant environments to enable students feel themselves free and have fun while learning English. Student 1 criticized the teacher creating an uncomfortable environment which hindered the student from feeling enjoyment and caused demotivation. In other words, Student 1 implied that enjoying the task at hand could mean enjoying the language learning activities and as a result an increase at motivation level. There were other students who had a parallel point of view.

> Extract 41. ... from time to time yes. After doing my homework some teachers congratulate us and this makes us happy (S5).

As Student 5 pointed out, positive feedback, which is a significant part of CD 3, from teachers could be a way of creating suitable environments in which students could feel free and enjoy their learning environments.

Students were also asked to what extend the activities used in classes contribute to their feeling relaxed and enjoying the language learning experiences for their own sake. They were also asked to state how well the activities get them to take their own responsibilities. Student 3 from experimental group made negative comment in this respect.

> Extract 42. Right now I do not feel in that way. Because there is a classical system and a book. Most of our classes based on the book and grammar, and rules are very important. But on the other hand, in speaking classes there are such kind of activities. If it was possible to apply similar things into grammar activities, it would be possible for me to learn how to learn (S3).

As can be seen, Student 3 thought that type of classes had different effects. When it was considered that speaking classes had been carried out by foreign teachers, it would not be wrong to assume that nonnative language teachers might have caused problems. Together with some other students, Student 3 indicated that nonnative main course teachers generally adopted a traditional approach in which

they mainly focused on teaching grammatical structures and which caused learners to be less autonomous. In the same vein, Student 7 from the control group highlighted the overuse of grammar teaching.

> Extract 43. ... there are some applications in speaking classes mostly. Apart from that not much, mainly grammar education (S7).

However, there were students who had similar views but who adopted different perspectives regarding the reasons underneath.

> Extract 44. I think our teachers try to be creative but students are unable to adapt themselves. They generally ask for traditional way of teaching and grammar education, and this causes teachers go that way ... (S6).

Student 6 was among the ones who thought that the reason for unpleasant atmosphere and being exposed to grammar instruction to much was not primarily the teachers but the students who demanded this kind of classrooms. Maybe this could be due to the existence of many students with high avoidance motivation.

There were students in both groups who believed that the language education they were taking was successful enough to make them independent learners.

> Extract 45. ... in general I learn how to learn (S9).

Some students like Student 9 from the control group stated that activities carried out in classrooms helped them become independent learners.

> Extract 46. Yes. ... there is an online platform and before I was not aware of the existence of such kind of systems. There were only books. I understand that online education is more active and productive. It is possible to see the use of creative applications especially in some teachers' classes (S4).

Although some students believed that activities used by teachers did not work out in terms of their capability to make them independent, Student 4 from experimental group expressed a different point of view. This student acknowledged that depending merely on coursebooks in a lesson was not so enjoyable; this student found the use of digital sources and platforms accompanied by the coursebook effective and implied that this was a good way to improve students' autonomies.

CD 4: Ownership and Possession. When it comes to evaluating the language education in terms of CD 4: Ownership and Possession, the biggest difference between the groups was observed. The mean score of experimental

groups was 5.8 whereas it was 3.8 for control group. Students were asked to evaluate the language education in terms of to what extend it helped them feel that was their program and they felt they were a part of the program. Students from experimental group indicated that they were feeling themselves as a part of language preparation program.

> Extract 47. I can use the phrase 'this is my program' in sentences ... (S3).
>
> Extract 48. ... I feel like this is my place in a way (S2).

Student 3 and Student 2 from the experimental group directly stated that they were feeling the sense of ownership and could define the program as their own. Another student from the experimental group approached the issue with a slightly different perspective and stated that:

> Extract 49. Yes. Our relationships between our teachers and friends make us feel that we are a part of this school (S5).

As can be understood from the above extract, Student 5 connected CD 4: Ownership and Possession with CD 5: Social Influence and Relatedness. It was evident that positive relationships between teachers and students, and students and students contributed to the construction of the feeling of ownership. Although students from the experimental group positively commented on CD 4, a student from this group had a different view.

> Extract 50. There is nothing related to ownership. Everybody just wants to go to their departments in a way. I don't think that nobody wants to stay here voluntarily. If they got a chance, they would leave the program and go and try to learn English from other courses (S1).

It could be said that Student 1 from the experimental group disapproved the compulsory nature of language preparatory program. Similar point of view could be seen among the students from the control group.

> Extract 51. I don't feel I belong to here (S8).

Some students from the control group directly opposed the idea that they had the feeling of ownership.

> Extract 52. ... not many things are done to feel ourselves a part of this school and I feel that I belong to my department more. Here in this school it seems that students should go immediately (S6).

Some other students like Student 6 stated that the time allocated for foreign language preparation program would not be enough to enable to develop such

feelings toward the program. Indeed, all students could take language preparation program for only one year and it seemed that this was not adequate. Maybe, such a big difference between students' views regarding this CD could be a result of this.

CD 5: Social Influence and Relatedness. In the following question of the interview, students were asked what they were thinking about learning in a social environment together with their friends. Generally, students in both groups agreed that it was important to be in social learning environment for sustainable goals.

> Extract 53. When I realized that I cannot speak English, I work harder. But although this helps me get good grades from the exams, it is very difficult to survive in real life as an individual without my friends' help (S1).

Student 1 from the experimental group indicated that individual study could help the students in a limited way. Basically, a student could get higher marks from exams individually but for long-term goals required using the target language in real life, which meant being a fluent speaker. In this respect, students could get help from their friends and practice their language with them. That means in nonnative environments in which students get limited opportunities to practice their language skills with native speakers of the target language, peer interaction becomes very critical. There are other students who thought in the same way in control group as well.

> Extract 54. I know when I am with my friends and study with them, I learn better (S7).

Therefore, it could be said that students from the both groups believed that they perform better when they work with their friends. Then, students were asked to what extend a social learning environment existed in their classroom. Almost all of the interviewed students stated that there was a partial interaction in the classroom and mostly they did individual study.

> Extract 55. My learning style has always been individual. And here at this school, I think my individual study make me go on (S2).
>
> Extract 56. I think it is mostly about 80% is individual (S3).
>
> Extract 57. ... only individual activities ... (S7).

As can be understood from the above extracts, Student 2 and 3 from the experimental and Student 7 from the control group directly stated that they preferred studying alone. There were students with similar views who thought that they rarely took part in interactional activities when their coursebook required or when they

were in speaking classes. Indeed, this finding was consistent with the previous findings which showed that mostly grammar based, traditional way of language teaching and learning activities which require students to be passive learners instead of active ones who get into interaction with their peers took a significant part of the education.

> Extract 58. When we do group work it is very good. In speaking classes, we do such kind of things (S9).

Student 9 from the control group pointed out that working with friends on a project was enjoyable, although in main course classes or with their nonnative teachers, there was no space for such kind of activities.

> Extract 59. We have project assignments in speaking classes. In general, except from speaking classes, students are diverted into individualized studies (S6).

Similarly, Student 6 agreed that speaking classes made them learn together as a group. But there were also students who thought that main course teachers also made use of group works in classes.

> Extract 60. Yes. For example, in our coursebooks there are activities that require pair work. But apart from the coursebook, it is rare. Some teachers have given group or pair work once or twice. Mostly through the coursebook (S5).

Participants like Student 5 indicated that pair or group works existed in their classes, but not because teachers wanted but because they were suggested by their coursebooks. In addition, Student 5 stated that some teachers had assigned these kinds of tasks once or twice. When the size of the course they were doing (720 hours) was considered, these numbers meant nothing.

The mean score for the experimental group was lower than the previous CDs discussed so far, which was 3.8. and the participants in the control group were given slightly higher points for this CD and the mean score for this group was 4.4.

CD 6: Scarcity and Impatience. Students were asked to evaluate the language learning environments in their classrooms and the activities used throughout language learning experiences in terms of the existence of scarce and valuable things. Students from both groups were given the lowest points for this CD and there was not a considerable difference between the groups. The mean score for experimental group was 2.8 whereas it was 2.4 for the control group. In the first

part of the interview, students discussed the existence of scarce and valuable things in their classroom.

> Extract 61. People who try to speak in English. They are very scarce. I mean both teachers and students. In addition, the funny times in classroom is scarce. Sometimes we play games and there are usually enjoyable but very few (S1).

Student 1 from the experimental group stated that what was scarce in the classroom was the use of language, which was expected to be the most prevalent thing. The criticism made by Student 1 was not only toward the other students' approaches but also for the teachers'. Normally, in language classes teachers are expected to use the target language as much as possible unless the method or the activity requires otherwise. But as could be understood from the above extract it was not the case. Moreover, it was implied that other students were not in favor of improving their language skills by practice. This finding was in line with the aforementioned findings of previous CDs. Thus, this could be counted as another evident for the presence of traditional way of language teaching and learning. Another issue that was brought up by Student 1 was the lack of humor which caused rather dull environment for learning a foreign language.

There were students who had a different approach.

> Extract 62. Today we can find English as a foreign language education everywhere. Here the education is free, so maybe in this sense I may not find any other free education opportunity outside (S2).

Student 2 from the experimental group focused on the economic side of the educational services provided by the school. For this student, what was valuable was financial benefits that were offered by the program. Different from these two students, there were students with opposing ideas.

> Extract 63. I don't think there are things that make me struggle to get them. But for example, in my department, ..., there is a lab which only can be used by students who take part in projects and you really should show that you deserve that privilege. ... There is nothing like this here (S6).

Student 6 from the control group thought that there were not any applications or uses of scarce things. This student stated that limiting the accessibility of a facility could add a value and turn that facility into a status symbol, which is a clear example of the use of CD 6. However, the same student complained about the deficiency of

such applications in school of foreign languages and implied a decrease in motivation.

According to the students participated into the interviews, among the other displays of scarcity were the use of time restrictions and prizes. Some students stated that teachers used prizes such as positive feedback for satisfying answers or task completion, or small tangible badges for each successfully completed assignment.

> Extract 64. Sometimes teachers give positive oral feedback to the ones who do well, that's all (S5).

As can be understood from the above extract, Student 5 from the experimental group found the rare use of positive oral feedback valuable.

Student 1 from the experimental group found the use such things stimulating. However, Student 3 from the same classroom did not approve this and stated that:

> Extract 65. I think that elements like prizes shouldn't be used in education. I think, they result in more individualized education. the fewer those special prizes, the better. When are used at individual base, they case some students drop out (S3).

What Student 3 implied was a critical point. This student acknowledged that using prizes worked out and had a kind of energizing power, but not in the long run. That is, at first, use of prizes could help students get motivated and take action, but as time passed and when some students realized that they were behind their classmates and prizes became unreachable, then the motivating power of prizes became inadequate and might have caused students to stop studying. There were students in both groups who thought that the use of prizes did not have an effect on their motivation.

> Extract 66. ... they are not valuable (S4).
>
> Extract 67. ... for me the only prize is to be able to learn English at the end of the year (S1).

Student 4 stated that the prizes did not have any effect on him/her because they were worthless. And another student, Student 1, pointed out that what was important was the goal of learning English and this could be interpreted that the use of prizes might not have any effect in this sense.

CD 7: Unpredictability and Curiosity. The mean score for experimental group was 4 whereas as it was 3.4 for control group, which indicated the existence

of a slight difference. In relation to this CD, many students thought that it was not possible to see things that make them curious or things that were new.

> Extract 68. ... I don't think that there are many things that make me curious. Mostly routine and not much new or interesting things (S3).

As stated in the previous section, Student 3 criticized the monotonous atmosphere of the classes and indicated that lack of elements which could catch students' attentions or which were interesting enough to trigger students' motivation caused them to feel demotivated.

> Extract 69. Actually, it is possible to make the topic interesting ..., which we never do. Generally, everything is thing is routine, nothing new (S1).

Student 1 from the experimental group stated that although it was possible for teachers to integrate interesting and novel things into the activities during classes, they did not prefer to do so and continued with conventional ways of teaching.

> Extract 70. ... usually there are routines, not much new or interesting things (S4).
>
> Extract 71 ... only grammar ... (S7).

In similar way, Students 4 from the experimental and Student 7 from the control group addressed the widespread use of conventional language teaching and learning activities and scarcity of new and motivating elements.

CD 8: Loss and Avoidance. Finally, towards the end of third step of the pre-intervention interview, students were asked to evaluate their language learning classes in terms of the existence of avoidance motivation. For this core drive, CD 8: Loss and Avoidance, there was no difference between the two groups whose mean scores were both 5.2, which was quite high. Besides, this was the only CD that all students both from the experimental and control groups anonymously agreed upon.

> Extract 72. ... because it is not about this school but our past education. This was always like this, it was enough to pass 50 ... And this is what happens in this classroom, and a lot. My friends do not study and their aims are not to fail (S1).

Student 1 from the experimental group stated that avoidance was a big problem and this was an issue that they brought with themselves to this school. Students were trying not to fail instead up striving for their meaningful goals.

> Extract 73. Yes. There are students who come to the classroom just not to be dismissed due to their attendance problems. There are also students who don't do their best and content with less (S3).

Similarly, Student 3 from indicated that some students lost their aims and just tried to fulfill some basic requirements such as attendance limits. For this kind of students, it might be possible that they tried to manage to go on their education at the lowest achievement level because they probably had difficulties in following the classes which caused attendance problems.

> Extract 74. The education here supports students' avoidance motivation a lot. One of our teachers counts everything in the classroom and take attendance by counting minutes, and decreases our grades when we are unable to complete tasks. So, we just try to complete those tasks. I know that if I don't get late and do my homework in a way, then it is not important whether I learn or not (S6).

According to the comment made by Student 6 from the experimental group, teachers' attitudes and applications had great impact on students' motivation. It could be interpreted that if a teacher does the right things in classroom environment, then it could be possible to foster students' motivation in the desired direction and could make students work for meaningful goals. Otherwise, wrong applications and attitudes could canalize students studying to avoid failure itself. Above extract exemplified this situation. A teacher in control group did not focus on the performances of the students but instead on the quantifiable data. According to Student 6, this caused students to feel demotivated.

Having completed the interviews, students were asked to evaluate the language skills which they thought to be problematic in their classrooms. Many students indicated that their teachers do not talk in English and this had negative effect on their desire to communicate in English. On the other hand, they also indicated that speaking classes were the ones in which they enjoyed participating. In this respect this skill might not be such problematic. Secondly, as could be understood from the above extracts, they strictly follow their coursebooks and are exposed to grammar classes too much. In addition, reading was not problematic for them as well. However, as to the writing activities, all students agreed that their classes lacked the proper use of writing activities. Although some teachers assigned writing tasks in line with the coursebooks, they did not try to make those activities enjoyable and motivating.

The grades given by the students and teachers for each dimension are provided in Table 73. To enable better comparison, survey results that are given in previous table can also be given in this table, too.

Table 73

Octalysis Scores from Pre-Intervention Semi-Structures Interviews and Survey

	Experimental Group				Control Group			
CD	Students	Teachers	Survey	Mean	Students	Teachers	Survey	Mean
Epic Meaning and Calling	4.6	3	3.84	3.81	4.6	3	3.83	3.81
Development and Accomplishment	4.8	4	6.03	4.94	4.8	4	6.08	4.96
Empowerment of Creativity and Feedback	4.6	4	5.53	4.71	4.4	4	5.27	4.56
Ownership and Possession	5.8	5	3.70	4.83	3.8	4	3.72	3.84
Social Influence and Relatedness	3.8	3	6.58	4.46	4.4	5	6.23	5.21
Scarcity and Impatience	2.8	6	6.36	5.05	2.4	5	6.32	4.57
Unpredictability and Curiosity	4	3	5.51	4.17	3.4	4	5.76	4.39
Loss and Avoidance	5.2	6	5.22	5.47	5.2	7	5.44	5.88

Before calculating the Octalysis score for experimental and control groups, to obtain more comprehensive results, the survey sores and interview scores of students and teachers were compared and the average score for each core drive in both groups was calculated. The Octalysis scores which are calculated by summing the squares of average scores for each CD was 177.14 for experimental group whereas it was 176.53 for control group. As can be seen from Table 73, there were not many remarkable differences between the experimental and control groups before starting the gamification intervention. The noteworthy differences were on core drive Ownership and Possession (Experimental Group: 4.83, Control Group: 3.84) and Social Influence and Relatedness (Experimental Group: 4.46, Control Group: 5.21). As a result of the above findings, the Octalysis scores for the two groups were computed and the Octalysis graphs, which represent the general situation in experimental and control groups, which are shown in Figure 15 and Figure 16, were drawn.

Octalysis graphs for general evaluation

Octalysis framework provides a visual tool that depicts the factors underlying motivation (Chou, 2015). It is also possible to detect whether long-term or short-

term motivation exists in the system. If the white hat CDs (CDs on the upper side of Octalysis framework) are more powerful, then it means it may be possible to achieve long-term motivation. Otherwise, black hat CDs (CDs at the bottom of the Octalysis framework) would result in short-term and unsustainable motivation. Another advantage of visual representation of Octalysis is that it can be possible to find out whether students are intrinsically or extrinsically motivated. If the CDs on the right side of the Octalysis are stronger, it means that students are intrinsically motivated, which is a desired situation. However, having stronger CDs on the left margin indicates the existence of extrinsic motivation.

One of the critical points to keep in mind while studying the Octalysis framework is that there are not good or bad CDs. That is, it is not expected to set the black hat CDs or right brain CDs to zero. What is important is to utilize each CD to make a transition to long-term motivation. In other words, it is important to able to diagnose what kind of role each CD plays in terms of motivation. There will be times when black hat CDs will be compulsory. Black CDs could be utilized as triggers for short-term motivation. Another issue is that in scope of this study only 1st level application of the Octalysis was carried out. There are four levels of Octalysis at each of which attractor states of students and learning styles are evaluated and used for more comprehensive system designs.

Figure 15. Pre-intervention Octalysis graph for experimental group, general evaluation.

Figure 16. Pre-intervention Octalysis graph for control group, general evaluation.

Although it was possible to mention the existence of a balance framework for each group, the biggest problem in both groups was the strength of CD 8: Loss and Avoidance and the weakness of CD 1: Epic Meaning and Calling. This showed that students in these classes had short-term motivation rather than more sustainable goals. In this respect, when the Octalysis for experimental group is investigated it can be stated that there was a balance in the distribution of the core drives. Before applying the gamification design, both approach and avoidance motivational drives were active. It was possible to say that students were studying more not to be unsuccessful rather than to achieve bigger and meaningful goals.

In addition to the general evaluation of language classes, students and teachers of both groups had been asked to diagnose the most problematic issues they had been experiencing. As result, it was found out that writing skill came forward. Therefore, pre-intervention interviews findings in terms of writing evaluation were investigated in relation to the eight CDs of Octalysis.

Pre-Intervention Teacher Interviews, Writing Evaluation

When interview findings were examined, it was found out that not all teachers were responsible for carrying out writing tasks in classrooms. In both groups, teachers had shared the responsibilities in terms of following specific activities. That is, one teacher was totally responsible for the writing tasks and the other teacher/s in that classroom did not interfere into writing processes. Therefore, only one teacher talked about writing processes in each group.

CD 1: Epic Meaning and Calling. In terms of writing activities in experimental group, the teacher stated that writing tasks were being regularly assigned at the time of the interview. In addition, it was understood from extract below, the teacher gave importance to whether these tasks help students follow larger goals. From the extract below, it is seen that the teacher had a bottom-up approach in writing. This might be due to the general pacing of the coursebook or teacher's own choice. Another issue that was not clear in interview was that whether the goals set by the teacher were long-term or short-term. For the experimental classroom teacher, the magnitude of writing evaluation on Octalysis was 4 on the 1-8 scale.

> Extract 75. I assign writing tasks a lot. Previously, they were just about writing short paragraphs. And I can say that students are following my instructions. And I try to set up goals for them (T1).

The situation in control group was not much different in control group, either.

> Extract 76. I don't think that there were meaningful goals in my writing classes. But I can say that students have a goal somewhat. Generally, they write just to complete the task and I frequently see papers that have been copied somewhere else or unrelated to the task (T3).

In control group, the teacher stated that writing activities in the classroom made students have goals in a way. But the above extract evidently shows that in writing tasks not only CD 1 was in play, but there were also other CDs such as CD 8: Loss and Avoidance that had been affecting the goal pursuit. Thus, it could be stated that instead of meaningful long-term goals, students had short-term ones which were the result of CD 8. This teacher gave 4 for the power of writing activities in control group.

CD 2: Development and Accomplishment. In terms of the second core drive, the teacher from the experimental group stated that it was possible to feel that students were developing their writing skills. However, when the teacher was asked to evaluate the situation from students' points of views, the teacher was not certain about that. This finding can be interpreted in the way that teacher was satisfied with the way he/she was conducting the writing activities whereas it was not possible to talk about the existence of highly motivated students. In addition, the teacher in the experimental group did not comment on the number of students who had been improving their writing abilities.

> Extract 77. Yes. I can feel that they are improving their writing, but ... do they feel the same ... I am not so sure (T1).

There was a similar situation in control group, too. The teacher of control group stated that somehow it was possible to see that students were improving their writing abilities thanks to the activities carried out.

> Extract 78. I think they are improving. I have accumulated some initial works of students that they wrote in the first month of the academic year. When I compare their current writings to those, there is difference. But, this is not for all students (T3).

In control group, based on the previous studies of students, the teacher argued that it was possible to see that students were improving. But as can be understood from the above extract, it was limited to some students and was not true for all students in the classroom. Moreover, like the teacher of experimental group, this teacher also tried to evaluate the situation from personal point of view. It was not possible to reach conclusions regarding the motivational levels of students or how the students feel about their progress from teachers' perspectives. Therefore, when above findings were put together, it might be possible to state that the teachers who were responsible for writing activities were not so interested in how their students feel about their own developments. On the other hand, in both groups, the existence of demotivated students could indicate a problem. The average of the grades given by the teachers for this CD was 4 (Experimental group: 5; Control Group: 3).

CD 3: Empowerment of Creativity and Feedback. Teacher from the experimental group said that there were one or two students who had been autonomous and who regularly wrote and ask for feedback.

> Extract 79. Yes, there are one or two students. They complete every task and come and show me their papers. These students are also curious about my feedback. But for others, I cannot say the same things (T1).

But as in the previous CD, the number of autonomous students was limited to one or two students, which barely composed 10% of the total number of students. Above extract shows that not many students were eager to write and ready to take their own responsibilities. In addition, the feedback mechanisms for writing tasks were given only when students demanded, which could be a sign of inadequate feedback.

> Extract 80. I have some students who really struggle to write. I encourage them to write at home and say that to write something, they don't need to wait an assignment to be given. But the number of students who do this is very low. The rest either don't write or had in incomplete tasks (T3).

Like experimental group, the teacher in the control group talked about similar issues. Accordingly, the number of students who had become autonomous enough to write independently was very low and mainly the majority of the students were indifferent to writing. Moreover, the teacher of the control group did not talk about any feedback mechanisms. The average grade given by the teachers for this CD was 4 (Experimental group: 4; Control Group: 4).

CD 4: Ownership and Possession. When the teacher of the experimental group was asked to state what he/she thought about the effect of CD 4, the teacher told that:

> Extract 81. That's very weak ... they are not in the position of owning their writings (T1).

A similar comment was made by the teacher of control group.

> Extract 82. I don't think that students could produce writings that are totally their own (T3).

So, according to the findings about this CD, the power of CD 4 was lower than the first three ones. The average for this CD was 3 (Experimental group: 3; Control Group: 2).

CD 5: Social Influence and Relatedness. In terms of how effective the social dimension on students' writing activities, the teacher of the experimental group said that usually students got help from each other.

> Extract 83. since they cannot write on their own independently, they certainly get or seek help from their friends at word level and so on. Sometimes they ask the structures that are supposed to be used in the task to their friends. So I can say that they write together (T1).

As can be understood from above extract, in the experimental group, student interactions with each other were limited to in task activities. And probably weaker students look for help from stronger ones in terms of writing ability because the teacher said that some students did not find the proper vocabulary and needed help. In addition, another part for which students sought help for was the grammar dimension of writing. However, there was not any indication of other kind of interactions in writing activities or group works in which students could learn from

each other not only at vocabulary or grammar levels but also in other ways. The teacher who was responsible for the writing activities in the control group had a slightly different view, though.

> Extract 84. I try to force my students work together on writing tasks. I either pair them or group them. But most of the time it ended up in an unwanted way for me. What I see is that weak students prefer to stay passive and refuse to participate. This puts pressure on good students. They complete the tasks alone not to get low marks but tell me it is the group product. But I know it isn't (T3).

Similarly, the teacher in the control group acknowledged that there was interaction and group work between students to some extent. However, the teacher added that the real situation was quite different. So, this could be interpreted in the way that a limited and superficial interaction existed in this classroom. The average grade given to this CD by the teachers was 4 (Experimental group: 4; Control Group: 4).

CD 6: Scarcity and Impatience. The teacher in the experimental group did not comment on this drive by stating that there was no application of any scarcity in the classroom for writing activities and gave one for this CD. On the other hand, the teacher in control group stated that she frequently offered extra marks for writing tasks completed in time and for tasks that were outstanding.

> Extract 85. I usually try to motivate my students by offering them extra points. If they could hand their tasks in before due time I give extra. In addition, I distribute paper badges for successful task completion. I think this works ... (T3).

Contrary to the teacher of the experimental group, the teacher of the control group stated that he/she made use of scarcity in the classroom. The teacher used points and badges which were among the most frequently used prizes and limited the time by use of due time. Therefore, it could be said that the teacher in this group merged CD 6: Scarcity and Impatience, CD 4: Ownership and Possession, and CD 2: Development and Accomplishment by utilizing points and badges. Because collecting badges could make students feel development, and improve their sense of ownership. Consequently, students could strive to protect their rewards and go on writing.

As expected, the magnitude of CD 6 in the control group was greater than the experimental group, but the average was 3 (Experimental group: 1; Control Group: 4).

CD 7: Unpredictability and Curiosity. In terms of curiosity, the teacher of the experimental group indicated that there were some students who were curious and seek new information.

> Extract 86. I can say that there are really curious students (T1).

Although the teacher stated that some students had a curious personality, he/she did not talk about the nature of the activities utilized in writing classes which had the potential of triggering students' curiosity and leading them to seeking for new knowledge. In this sense, it could be stated that there were not many interesting and/or new writing tasks for students in the experimental group.

The control group, the teacher had a similar point of view.

> Extract 87. Yes, I try to find interesting topics for children and assign them different topics. But curious students are the ones who tend to like them (T3).

In the control group, the teacher admitted he/she paid attention to turn routine writing tasks into interesting ones but also added that some students were not curious and there was no difference for them in terms trying to utilize interesting topics in the classroom. The average for this CD was 5.

CD 8: Loss and Avoidance. For the last CD, the teacher of the experimental group stated that both approach and avoidance motivation could be seen in the classroom. What was important in this classroom was the attitude taken by the teacher.

> Extract 88. When I force them to complete the task that I have assigned, then all the students do it. But I don't force them. So, nearly 9 students do always complete their tasks whereas the others do not give any papers at all. I can say that students write not to contradict to me (T1).

When the teacher made it compulsory and indicated that there would be penalties for incomplete writing tasks, more than half of the students would have handed in their works. The teacher did not prefer this approach and utilized voluntary participation. This might have resulted in an increase at motivation levels of some students and triggered long-term motivational goals. However, the higher number of demotivated students who did not take part in writing activities might be

a problem even though the teacher managed to motivate some students. In the control group, the teacher indicated that avoidance was a big problem.

> Extract 89. As I said earlier, some of my students usually prefer to stay silent and just write their names on the final paper just not to lose marks. I think this is a sign of avoiding. In addition, some other students hand in tasks that are carelessly completed. Again, not to get low marks (T3).

According to the teacher of the control group, avoidance motivation was very high in the classroom. As can be understood from the above extract, many of the students completed their tasks just to avoid negative consequences such as low marks. The average grade for this CD was 6 (Experimental group: 6; Control Group: 7).

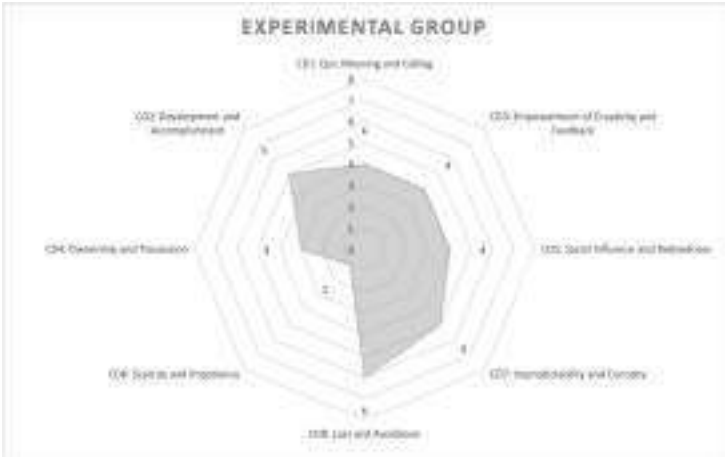


Figure 17. Pre-intervention Octalysis graph for experimental group, teachers' writing evaluation.

Figure 18. Pre-intervention Octalysis graph for control group, teachers' writing evaluation.

As can be seen in Figures 17 and 18, the grey areas that depict the formation of Octalysis represent a problematic collapse onto the bottom core drives (CD 6, CD

7, and CD 8). This finding indicates that students' motivation was not longstanding. Instead, many students were trying not to fail and they did not have control of their own behaviors. Moreover, these graphs could also be the indicators of students' being inclined to put off their duties in terms writing tasks, which could be understood from teachers' statements as well.

According to the teachers, the main difference between the experimental and control groups were about CD 4: Ownership and possession and CD 6: Scarcity and Impatience. In terms of CD 4, it could be said that students in the experimental group felt that they owned their works, which would probably be because they were more motivated to improve their writing skills and protect their successes. On the other hand, as Figures 17 and 18 suggested, it was more possible to see the utilization of scarcity techniques in the control group than the experimental group. This was also confirmed by the statements of the teacher, who said that he/she made use of points and badges, and limited the time while assigning writing tasks.

Apart from CD 4 and 6, there was a balance in terms of the distribution of other CDs on Octalysis framework. However, it should be noted that low power of CD 1: Epic Meaning and Calling could be considered as a major problem.

Pre-Intervention Student Interviews, Writing Evaluation

Having found out what the teachers in both groups thought about the eight CDs of Octalysis in terms of writing activities, in the next step how students had evaluated their writing classes was examined.

CD 1: Epic Meaning and Calling. When students were asked to evaluate their writing classes, there were two points of views that had been brought up. Nearly half of the students were thinking that the activities in writing classes helped them set meaningful goals.

> Extract 90. Yes, it offers, especially in terms of academic English (S1).
>
> Extract 91. ST 3 Yes. We write for meaningful goals. Because writing includes all the things (S3).

Although students 1 and 3 from the experimental group acknowledged that the activities carried out in the scope of the writing classes helped them develop meaningful goals, they did not directly give reasons for why they thought so. For Student 1, it was important to make progress in academic English, which could be

interpreted that academic knowledge might be a part of a bigger goal. Similarly, Student 3 stated to have meaningful goals. This student might feel that to be able to write well could be a sign of knowing a language.

There were two students in the control group who indicated that the activities were helping them to get closer to bigger goals. But they did not verbalize their thought. They only confirmed that such thing existed and then passed to the grading CD 1.

Even though there were four students who made positive statements, there were six students who fell on the contrary.

> Extract 92. Actually, we do not so much in writing. There is a portfolio but we do not use it frequently. And there are no meaningful goals (S4).
>
> Extract 93. Not really. I still think that I am incompetent in writing and I don't have big goals in terms of writing (S5).

In the experimental group, Students 4 and 5 stated that writing activities were not successful in terms of creating opportunities for meaningful and bigger goal setting. They also pointed out that for successful goal setting, it was important to feel the development, which was CD 2. Since Student 2 did not feel himself/herself developed enough, he/she was not able to create meaningful goals.

Similarly, there were students who thought in the same direction.

> Extract 94. ... they are unsuccessful, we do not have meaningful goals. Everything is superficial and there are no feedbacks (S6).
>
> Extract 95. No. avoidance is more dominant (S9).

Like students from the experimental group, students 6 and 9 from the control group pointed out that to achieve meaningful goals, there are other CDs that should be taken into consideration because as in the case of Student 6, lacking appropriate feedback mechanisms might hinder students' approach motivation. In addition, as in the case of Student 9, when a student perceives higher levels of avoidance motivation, then it would be difficult to have students strive for their big goals.

To sum up, students from both groups evaluated their writing classes in terms of CD 1: Epic Meaning and Calling in the same direction and magnitude. The average for the experimental group was 4 whereas it was 3.8 for the control group. This finding was also in line with the teachers' views who gave 4 in each classroom.

CD 2: Development and Accomplishment. When it comes to evaluating the classes according to the CD 2: Development and Accomplishment, students from both groups had different views.

> Extract 96. I think my writing is developing. Although there are activities that will make me feel I improve my English, due to the high number of students in the classroom, teacher cannot find enough time for each student (S2).

As can be seen from the above extract, Student 2 from the experimental group admitted that the activities carried out in writing classes helped to improve writing. However, the same student found it problematic owing to the insufficient time allocated for each student, which would result in less feedback from the teacher. In this respect, it could also be stated that CD 3: Empowerment of Creativity and Feedback had an important role in improving the quality of CD 2.

> Extract 97. It is not negative. Of course, there is a development but not very slowly. Not at desired speed (S3).

Student 3 from the experimental group made positive statement like Student 2. But this student found it slow and thus unsatisfactory.

> Extract 98. There are different types of writing and as you progress you learn these (S9).

Although Student 9 from the control group had similar point of view, he/she did not clearly say that the activities helped him/her develop. But as stated in the extract, becoming competent in different types of writing would be an indicator of development.

There were students who thought that writing activities were not successful in creating an atmosphere in terms of CD 2.

> Extract 99. ST 4 I don't write much so I don't expect such a thing. In other areas I can feel I am developing but not in writing (S4).

Student 4 from the experimental group stated that he/she was already demotivated in writing classes. Therefore, although he/she found the activities in other areas successful, it was not possible to mention any success for the energizing power of writing activities regarding CD 2.

> Extract 100. ... no feedback. It is not possible to see the development (S6).

Like Student 4, Student 6 from the control group found the writing activities unsuccessful in making students feel the development. The reason for this was the lack of enough feedback by the teacher.

The above findings showed that there was a balance between the groups in terms of CD 2. However, final average grading was 3.6 for the experimental group whereas it was 4.4 in the control group, which was a slight difference.

CD 3: Empowerment of Creativity and Feedback. The other CD students commented on was the last CD of white hat core drives that are placed on the upper part of the octagon. When students were asked to what extend the writing activities were successful in getting them to become independent writers, a balance could be observed between the groups.

> Extract 101. There is some kind of creativity. I continuously change the writing topics and we try to find creative ways ... (S2).
>
> Extract 102. ... there is a space for creativity (S9).
>
> Extract 103. ... sometimes I write at home independently (S6).

The extracts of the Student 2 from the experimental group and Student 9 from the control group showed that in both groups there were students who believed that writing activities allowed to be creative which could be a part of autonomy. In addition, this view was supported by Student 6 from the control group who stated that writing activities made him/her an independent writer. However, the number of students who thought otherwise was higher.

> Extract 104. There are normal standard things ... (S5).
>
> Extract 105. ... concrete standard topic and applications (S7).

Student 5 from the experimental group and Student 7 from the control group stated that there were monotonous activities in classes which led to uninteresting applications. As a result, it was not possible for them to enjoy their own learning experiments, which caused demotivation in terms of CD 3.

When the grades given by the students were examined, it was seen that they were in line with the teachers' evaluation. The average grade for experimental group was 4 while it was 4.2 for control group.

CD 4: Ownership and Possession. When students were asked to state their views in terms of CD 4, there was not a significant difference between the two groups. The average grade in experimental group was 5 and it was 5.4 in control group. When the interview transcripts were examined, it was seen that many students in control group felt stronger ownership than the ones in experimental group.

Extract 106. ... it is high ... (S3).

Extract 107. I feel like the final outcome is mine because when I write, I try to do my best (S5).

Extract 108. I feel I own my writings (S6).

As can be seen from the above extracts, in both groups there were students who thought that CD 4 was strong in their classes. Yet there were also students with contrary views.

> Extract 109. Since I am not so successful in writing, I cannot feel the ownership (S2).

As Student 2 indicated that it might not be possible to feel the ownership better when the feeling of development was low. Since Student 2 felt his/her writing skill was not improved enough, the perceived success was relatively low and this caused Student 2 to feel less ownership.

Although, students indicated high levels of perceived ownership, this finding was not in line with the one from teachers' interviews. The averages from students (Experimental group: 5; Control group: 5.4) was clearly lower than teachers' evaluation (Experimental group: 3; Control group: 2). Apparently, students thought that the writing activities were more successful in terms of CD 4 than their teachers.

CD 5: Social Influence and Relatedness. CD 5 was the weakest core drive in the experimental and the second weakest in the control group. Almost all the students in both groups had negative thoughts regarding the social dimension of writing activities.

> Extract 110. We sit and write alone (S1).
>
> Extract 111. ... group activities in writing tasks, that never happens. Always individually (S2).
>
> Extract 112. ... writing is an individual process. (S9).

The extracts taken from Students 1 and 2 from the experimental group and Student 9 from the control group showed that students thought they were supposed to write alone without any interaction with their friends.

> Extract 113. Once I remember that I write with my desk mate, but after that I haven't done it with somebody else so far. (S3).
>
> Extract 114. I think there have been two writing tasks so far which have required group work. (S7).

There were also students such as Student 3 from the experimental group and Student 7 from the control group who thought that although there were examples of pair or group work, those were too few to mention the existence of a true social and interactional dimension in writing classes.

Students from the experimental group gave 2.8 on the average for this CD whereas the ones in control group gave 2.6. However, these grades were lower than the teachers' who gave 4 for both groups.

CD 6: Scarcity and Impatience. Another CD on which students and teachers had different thoughts was CD 6. Students in experimental group thought that there were some samples of scarcity in their classes.

> Extract 115. Sometimes teacher applies some rules and require us to use certain structures or vocabulary. If we can do it, we get extra points as prize. But even there are not tangible prizes when the teacher gives us positive verbal feedback for our efforts, this makes us happy (S2).

Student 2, for example, pointed out that when the teacher limited the writing tasks by applying some rules and offered a prize for successful task completion, it became a kind of motivator. In addition, according to Student 2, usually it was possible to see CDs act in chunks. When feedbacks were used for specific purposes than they became more valuable and resulted in an increase in students' motivation.

> Extract 116. ... there are not any uses of scarcity (S5).

There were students with different views. Students like Student 5 thought that it was not possible to find any applications of scarcity in writing classes.

> Extract 117. ... only time limitation ... (S8).

In the control group the situation was similar. Student 8 stated that the teacher used time as a scarcity item by limiting it. In this way, it was possible for teacher to active students.

The average points in both groups were not very different from each other (Experimental group: 3; Control group: 2.4). But as in the previous CD, there was a difference between the students' views and teachers' point of view.

CD 7: Unpredictability and Curiosity. In terms of CD 7, there was a consensus between students from both groups that writing classes were moderately successful in creating curiosity.

Extract 118. There are not many things that make us curious in writing ... I wonder about the reaction of teacher ... whether I am successful or not (S2).

Student 2 from the experimental group stated that there were not many things that made him/her curious and led him/her to seeking for new knowledge except for the reactions of the teacher. It could be said that teacher's way of giving feedback could be a source of curiosity.

Extract 119. ... nothing interesting (S4).

Extract 120. ... since topics are not interesting, I frequently get bored (S7).

However, more students stated that writing classes were not successful and did not include any items that were interesting and/or new. And this could be a source of demotivation.

The Students from the experimental group gave 4.2 for CD 7 whereas students in the control group gave 4. In addition, this was in line with teachers' evaluations who gave 5 for experimental group and 4 for control group.

CD 8: Loss and Avoidance. For the final core drive, there was not significant differences between the experimental and control groups. However, this was the CD with the highest marks in both groups.

Extract 121. ... most of the students avoid criticism and strict attitudes of the teachers because the teacher can make really harsh comments (S1).

According to Student 1, there reasons that students did not want to get feedback might be related to the attitudes taken by the teachers. Strict approaches and wrong choice of classroom language could result in demotivation or increase students' avoidance motivation.

Extract 122. Generally I try to avoid making mistakes. And sometimes I just write to complete the task (S5).

As student 5 pointed out, sometimes students' motivation to do a task was not to make mistakes, which could result in an increase in avoidance and decrease in CD 1.

Extract 123. Especially in terms of assignments, we write just to complete the task (S9).

In a similar way, Student 9 in the control group pointed out that frequently short-term motivation to complete the writing assignments played a big role in writing classes.

Extract 124. This is very powerful. When we do not write, it affects our grades and so I write not to be unsuccessful and get low grades (S6).

As can be seen from the extracts, for many of the students in both classrooms, the forcing drive behind their behaviors was not the desire to achieve their meaningful goals but to avoid the failure itself.

The average grade for the experimental group was 5.6 whereas it was 5.4 for the control group. And this was in accordance with the teachers' evaluations (Experimental group: 6; Control group: 7) who thought that avoidance motivation was widespread in writing classes.

Octalysis graphs for writing evaluation

Having completed the examination of the transcripts obtained from the students in the pre-intervention interviews, Octalysis frameworks based on their evaluations were drawn.

Figure 19. Pre-Intervention Octalysis graph for experimental group, students' writing evaluation.

When the Octalysis frameworks in Figures 19 and 20 that show the situations in writing classes from students' perspectives at pre-intervention stage are examined, it could be seen that there was a disproportion in terms of the distribution of CDs on the octagon. Although it is usually expected to have more powerful CDs on the upper side of the octagon and on the right side of it, the CDs at the bottom and on the left were much more prevailing. Especially the lowest values on CD 5: Social influence and Relatedness and high values in CD 8: Loss and Avoidance were very problematical.

Figure 20. Pre-Intervention Octalysis graph for control group, students' writing evaluation.

On the other hand, although on many CDs, there was consistency between students' and teachers' evaluations, there were also differences on some CDs. So, to overcome this problem, the students' grades and the grades given by the teachers for each group were summed up and a new average value was calculated. Pre-intervention overall evaluation of writing classes was given in Table 74.

Table 74

Pre-Intervention Overall Evaluation of Writing Classes

	Experimental Group			Control Group		
CD	Students' Mean	Teacher's Grade	Average	Students' Mean	Teacher's Grade	Average
CD1	4	4	4	3,8	4	3,9
CD3	4	4	4	4,2	4	4,1
CD5	2,8	4	3,4	2,6	4	3,3
CD7	4,2	5	4,6	4	4	4
CD8	5,6	6	5,8	5,2	7	6,1
CD6	3	1	2	2,4	4	3,2
CD4	5	3	4	5,4	2	3,7
CD2	3,6	5	4,3	4,4	3	3,7

By using the data provided in Table 74, it was possible to compute overall Octalysis scores for each group. The Octalysis score for the experimental group was 136.85 whereas it was 133.13 for the control group. This finding indicates that there were not significant differences between the groups before the application of gamification. In accordance with Table 74, the Octalysis frameworks that show pre-intervention overall evaluation of writing classes were drawn.

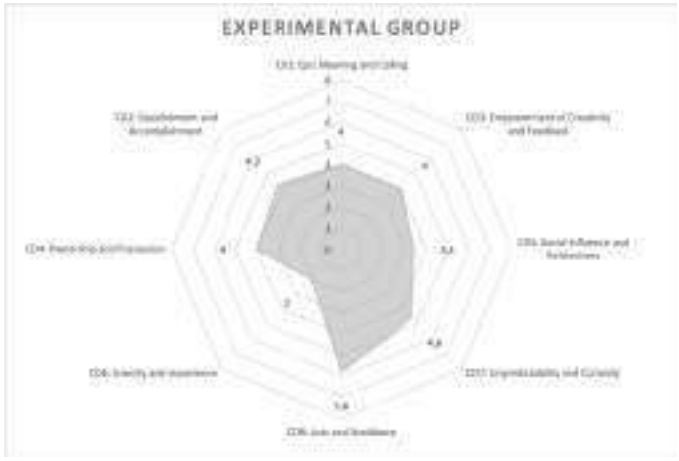


Figure 21. Pre-intervention overall evaluation of writing classes in experimental group.

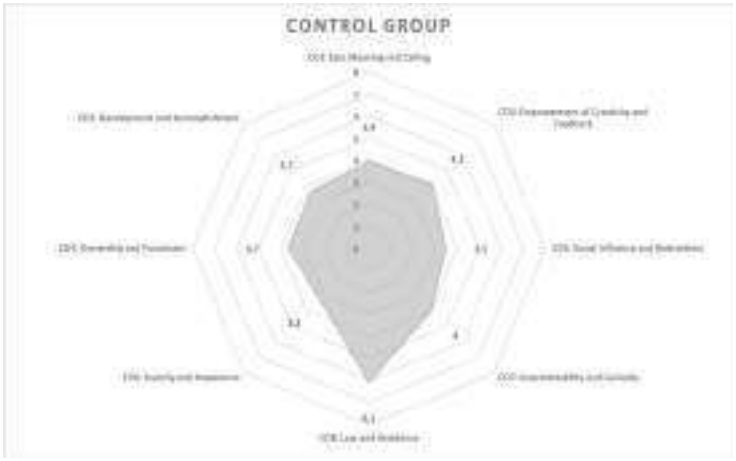


Figure 22. Pre-intervention overall evaluation of writing classes in control group.

Pre-intervention interviews with the teachers and students provided the data necessary to understand the weak and problematic sides of the language education in the school of foreign languages. At first step, it was found out that there were some problems that prevent students from getting motivated. In the next step, the students' and teachers' evaluations revealed that one of the most challenging part of the language teaching and learning activities was about the writing skill. Therefore, further examination of the interview data enabled the creation of the Octalysis frameworks which would be used as reference points for a gamification intervention for writing classes by taking into the weaknesses and strengths of the experimental group into consideration.

The biggest weakness was the high avoidance motivation and lack of social dimension. In addition, there must be a place for the activities that would contribute to the feeling of development and competence, and autonomy. In this way, it was

hoped to increase students' goal attainment motivation, work together, and enjoy their own learning experiences. In this respect, the first step was to draw the Octalysis Strategy Dashboard (Chou, 2015, p. 463).

Octalysis strategy dashboard.

After diagnosing the problems and understanding what kind of drives affected the motivation of students, in the next step it was time to have a plan. For successful applications of gamification interventions, it is important to know how to utilize critical elements to maximize the efficiency of applications. Chou (2015) suggests using another tool which is called as Octalysis Strategy Dashboard, which will help to understand the metrics, the users, and desired actions better. The strategy dashboard includes five important elements: business metrics, users, desired actions, feedback mechanics, and incentives.

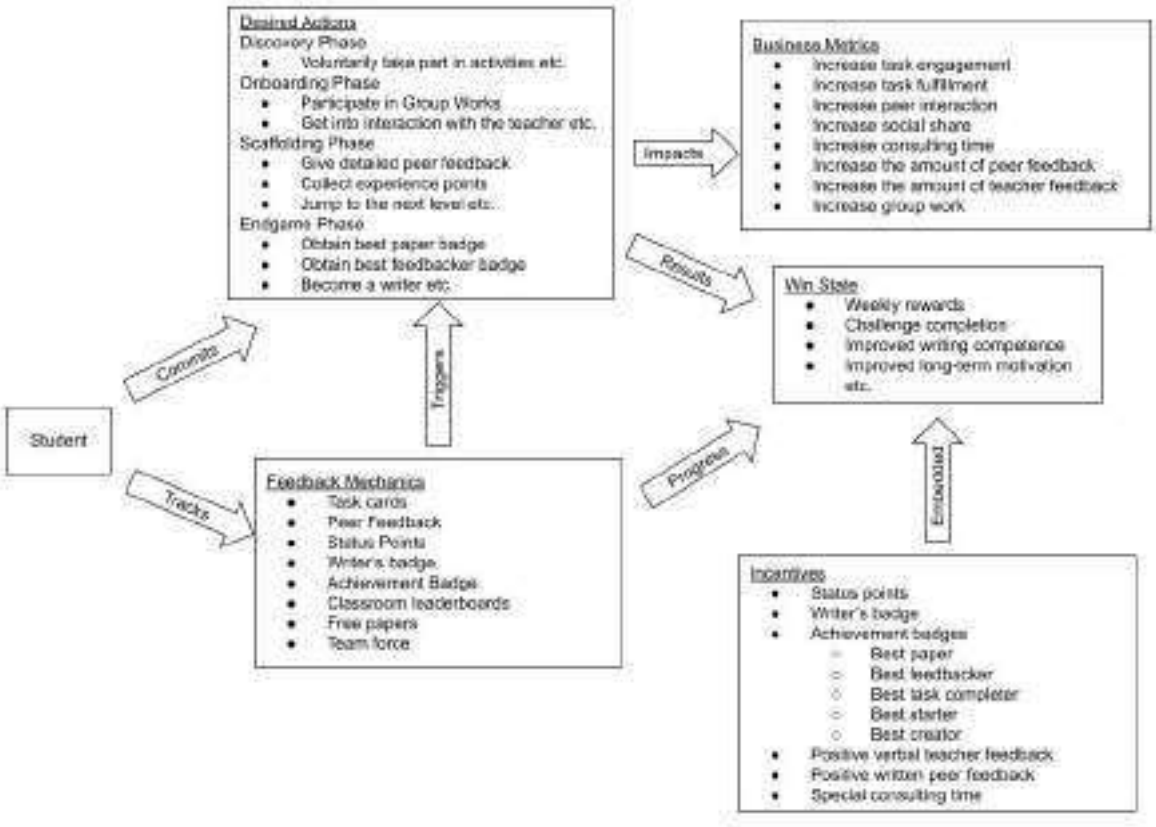


Figure 23. Octalysis Strategy Dashboard.

The player who was the students in current study commits the desired action. They can do this directly or by means of feedback mechanics which work as the

triggers to motivate students. This affects achieving the business metrics which are the goals and results in reaching the win-state which is aimed by feedback mechanics and supported by rewards.

Business metrics. Business metrics is a list of metrics that are measurable and lead to the game objectives. Indeed, these metrics consist of the ones which are required for successful projects. The metrics for the current study were:

- Increase task engagement
- Increase task fulfillment
- Increase peer interaction
- Increase social share
- Increase consulting time
- Increase the amount of peer feedback
- Increase the amount of teacher feedback
- Increase group work

In line with the finding of interviews and according to the resulting Octalysis framework, the above-mentioned metrics were defined. These metrics were all measurable ones, which meant any increase in these metrics could lead to successful gamification application.

The Octalysis framework for writing activities showed that CD 8: Loss and Avoidance was too high. Therefore, while determining the metrics, it was important no to increase this CD any more. The weakest CDs in experimental group's writing classes were CD 6: Scarcity and Impatience and CD 5: Social Influence and Relatedness. CD 6 which could be used to trigger some students' motivation. However, due to its being one of the black hat CDs and useful for short-term motivation, it should be cautiously integrated into the system but shouldn't be made one of the ultimate targets. On the other hand, CD 5 which was powerful in terms of its potential to increase the social interaction among students and in turn, support the development of CD 3: Empowerment of Creativity and Feedback was one of the primary point of focus. Therefore, in gamification intervention increasing task

engagement, peer interaction, group work and social share were identified as the major goals.

In addition, an increase in task engagement and task fulfillment was given importance since CD 2: Development and Accomplishment and CD 4: Ownership and Possession were not as high as desired in the experimental group. It was expected that as students engaged in writing activities more and as the number of their task accomplishment increased, their sense of development in writing competence and then feeling of ownership would also increase.

Finally, together with above-mentioned metrics which could result in an increase in CD 3: Empowerment of Creativity and Feedback, increasing the amount of feedback both by teachers and by peers was the other key goals of the design.

In the first study, the supportive roles of these CDS regarding the formation of DMC for English language learners had statistically been proven. Thus, it was expected that increases in the above-mentioned CDs would eventually result in an increase in terms of students' engagement in activities in the long-term.

Defining user types. For successful Octalysis gamification, it is crucial to know who the target users are. In the scope of the current study (indeed in all educational gamification), students and the teachers could be defined as the users who would be called as the players in gamification systems.

- Students in the experimental group who are going to take place in gamification design to improve their writing skills.
- Teachers who are going to use gamification to overcome problems in writing classes.

While planning Octaylsis designs, it is possible to define different desired actions for different types of players. In educational contexts, it is possible to say that player types are very similar to learner styles. That is, in educational environments, it is possible to define different gamification techniques to motivate students with different learning styles. In addition, in language learning classrooms it is possible to find students with varying degrees of writing competence. As the findings of interviews supported, there were students in the experimental group who believed they were good writers and very competent whereas there were also students who felt their writing skills were not so advanced. That is, there were

weaker and stronger students in the experimental group. Also, it should be noted that some students were already demotivated towards writing and some were more motivated than others, which might affect the application of gamification intervention. However, only level 1 Octalysis was utilized in the scope of this study. Since including students with different learning styles into the design can be carried out at level 2 Octalysis, this step was skipped and further planning was carried out in terms of desired actions.

Defining desired actions. In the next step, what kind of desired actions are expected should be written down. According to Chou (2015), the desired actions turn into win-states when they are accomplished by the students. The definition of desired action consists of four phases: Discovery phase, onboarding phases, scaffolding phase, and endgame phase. Chou (2015) makes such distinction because he states that the motivation levels cannot be the same throughout a game. The motivation at the beginning of a game and at the end are different. Indeed, this was an important part of Dörnyei et al.'s (2016) DMC theory. For this purpose, the actions and steps throughout the gamification process to be carried out by the learners were defined.

- Discovery phase
    - Learn about the intervention and give consent
    - Voluntarily take part in activities
- Onboarding phase
    - Participate in group works
    - Carry out paired tasks
    - Complete the task
    - Participate group discussions
    - Hand in independent free writing works
    - Get into interaction with the teacher
- Scaffolding phase
    - Complete brain storming sections of tasks

    - Organize suitable outlines
    - Give one holistic written peer-feedback each week
    - Give two holistic written peer-feedbacks twice in each week
    - Give detailed written peer-feedback
    - Collect as many experience points as possible in each week
    - Level up
- Endgame phase desired action
    - Obtain best paper badge
    - Obtain best task completion badge
    - Obtain best creative paper badge
    - Obtain best starter badge
    - Obtain best feedbacker badge
    - Become a writer

Defining feedback mechanics. To inform the students about their progresses and show that they are carrying out meaningful actions, feedback mechanics are used. Thanks to feedback mechanics, which act as triggers, students could measure their improvements. The first step of defining feedback mechanics is to decide the channels of interaction and communication, which were office hours, classroom notice board, classroom leaderboard, written papers by students, peer-feedback papers, teacher feedback papers, and classroom social media group.

The next step includes to placement of feedback mechanics. In scope of the current study, several mechanics were designed with the inclusion of following elements: Due date announcements (CD 6), task cards (CDs 7, 5 and 6), peer feedback (CDs 5, 3 and 2), status points (CDs 2, 4 and 6), writer's badges (CDs 2, 4, 6, and 7), achievement badges (CDs 2, 4, 1, and 7) class leaderboards (CDs 1, 2, 6, 7), free papers (CDs 3, 1, and 2), team force (CDs 5, 1, and 3). This mechanics were supposed to trigger one CD or a cluster of CDs in line with the desired actions.

Due announcements. The element 'due date announcement' was supposed to activate CD 6: Scarcity and Impatience because it put a limit on the time.

However, ‘due date announcements’ had the potential of activating CD 8: Loss and Avoidance. Therefore, to avoid over-use of CD 8 which was already a problem, students were assured that they would not face any penalties whatever their choices.

Task cards. The other mechanic ‘task cards’ was to activate CD 5: Social Influence and Relatedness, and CD 3: Empowerment of Creativity and Feedback. Each task card included a different special challenge, complete definition of the task and experience point allocated for that task. Since these tasks required to work in pairs and/or groups they were aimed at increasing the number of interactions in classroom. In addition, students could write down their own cards and create their own tasks, which was an important trigger for autonomous actions. Finally, the task cards were offered in closed format, which meant students could not see what is written inside until he/she chose the card. The goal was to make use of the power of CD 7: Unpredictability and Curiosity.

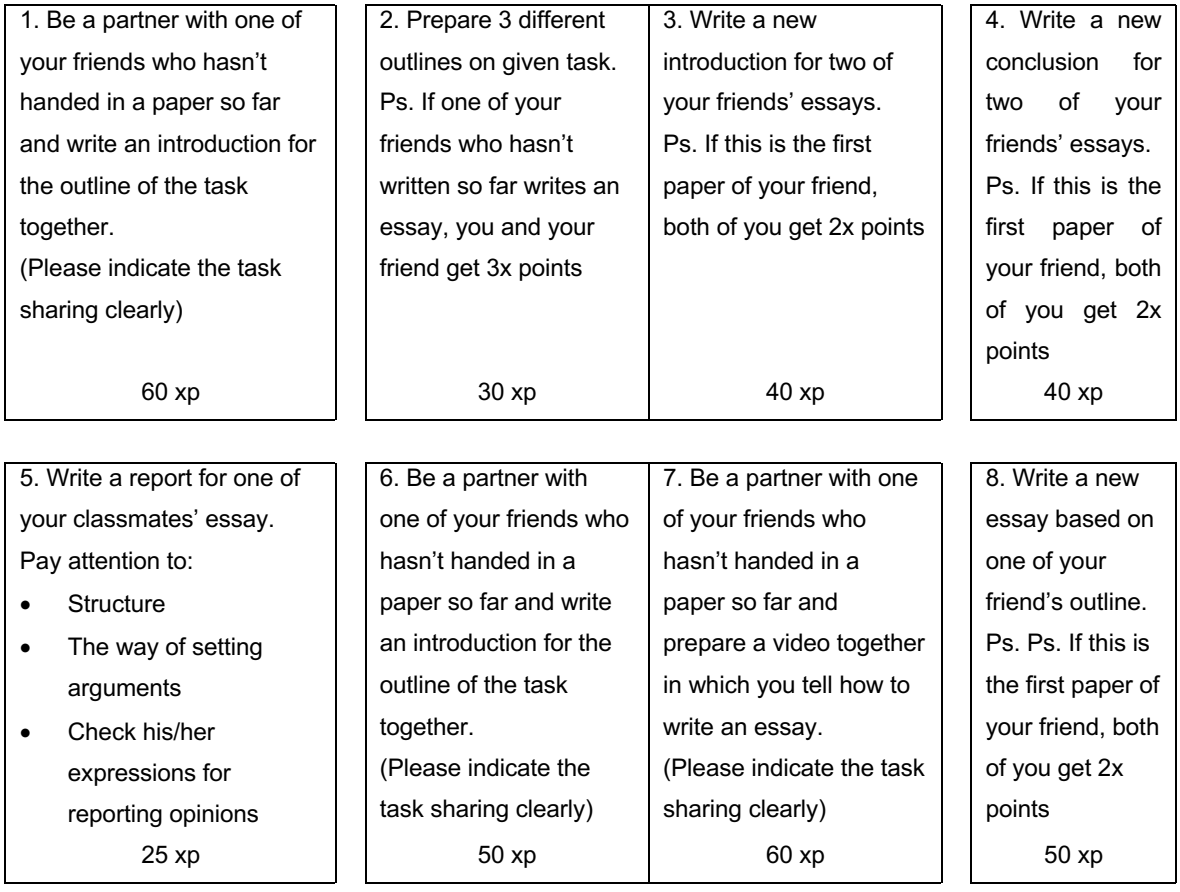


Figure 24. Sample task cards.

As can be seen in Figure 24, almost all the task cards 1, 6 and 7 included pair work. However, although the other cards seemed individual works, they contained tasks that required interaction with others.

Peer feedback papers. Another element aiming at increasing engagement in classroom is the use of peer feedback papers. 'Peer feedback papers' were especially integrated into the activities, because it was a worthy opportunity to trigger social interactions between students, which had been one of the most criticized aspects of writing classes both by the teachers and the students. So, it can be said that peer feedback papers were strengthening the effect of CD 5: Social Influence and Relatedness. Moreover, it was aimed to get students to take their own learning responsibilities by increasing auto-control. By utilizing CD 3: Empowerment of Creativity and Feedback in this way, it was possible not only to get stronger students help weaker ones but also help students see their errors without teacher pressure. Another use of this element was that students would observe their improvements as the number of positive feedbacks increase.

Status points. To improve students' feeling of development, the element 'status points' was utilized. For each successful task and/or sub-task completion a pre-determined experience point was given. In addition, students were given extra points that would show their creativity and extra-curricular activities. As their status points increase, they could feel they were improving their competences, which was an effect of CD 2: Development and Accomplishment. In addition, the more points they had the more they would feel the ownership and try to protect their earned places, which indicated the existence of CD 4: Ownership and Possession. Finally, the limitation of points with successful task completion included CD 6: Scarcity and Impatience.

Total status points include points from task completion and points for sub-sections of the essays. Students got xp points for completing the tasks required by task cards. If students want to write the essay than for each step they got extra points. For successful completion of 'brain storming' they got 10 xp, while they got another 10 xp for planning, 50 xp for complete essays, 50 for first peer-feedback they gave, and finally 50 for second peer feedback they gave.

Total Status = brain storming (10 xp) + planning (10 xp) + essay completion (50 xp) + peer feedback 1 (50 xp) + peer feedback 2 (50 xp)

Writer's badges. Another element used in the study was 'writer's badges' which were the pin badges with pictures of famous literary figures, motivation badges with positive feedback words written on, and small metal pins. As can be seen there were three groups of badges. In each week, students with the highest points were prized with 'the writer' badge. This was the metal badges which were small but more valuable because they were scarce. In this respect, CD 2: Development and Accomplishment was aimed to be increased. And CD 6: Scarcity and Impatience was also in action. In addition, this would also trigger the sense of CD 4: Ownership and Possession and allow students grow their collection of badges.

The second group of badges, 'best paper' badges, which consisted of pictures of famous writers were given for best essays.

Samuel Beckett | Frida Kahlo | James Joyce | Oscar Wilde | William Shakespear

Figure 25. Sample Best Paper Badges

Since these badges were only attained by the winner of the week, CD 6: Scarcity and Impatience was very powerful with this mechanic. To elevate this, the number of categories which were given a badge was increased. The biggest problem was the risk of losing weaker students who had seen that there were no possibilities for them to beat their stronger friends. Therefore, each week 'best starter' badge was given for the students who had not took part in activities in the previous weeks. This way, the inclusion of weaker students was possible. In addition, there were students who were on the way of developing their writing abilities. These students were not competent enough to produce perfect papers to compete with their stronger peers. However, they could perform better in sub-tasks such as giving peer feedback, being creative ideas, or task completion. Thus, the badges 'best starter', 'best

feedbacker', 'best task-completer', and 'best creator' were integrated into the system. Actually, these were smaller than the 'writer's badges' and were easier to obtain.

Classroom leaderboard. Another element used was the 'classroom leaderboard'. Each weak a list of points students got were announced in descending order. Thus, CD 1: Epic Meaning and Calling, CD 2: Development and Accomplishment, CD 6: Scarcity and Impatience, and CD 7: Unpredictability and Curiosity were activated with this mechanic. The problem with this element was the same with the previous one. There was a risk of losing weaker students when they got disappointed. Therefore, four different categories were created. Students were upgraded to the next category provided they had collected enough experience points. Each student started with white papers. Then, when they accumulated 100 points, they were upgraded into green paper category. For blue paper category 200 points were required whereas for red paper category the requirement was 400 points. Finally, students with 700 points were upgraded into yellow category.

Each week, as students were grouped according to their level of competence, separate leaderboards were announced. In this way, it was possible to obtain more homogenous groups. And each group was evaluated separately and badges were given for each category in each group. In this way, every student would get a change to gain a prize.

An important finding of pre-intervention interviews was that writing activities were not suitable for independent and creative tasks. To overcome this problem, a special use of prize was offered for free papers. That is, students were free to bring their free writings into the classroom and share their experiences in a way they prefer. This was prized with positive verbal appraisal by the teacher. And students who brought free papers also got the privilege of extra office hours in which they could receive special consultancy from their teachers. With this element CD 1: Epic Meaning and Calling, CD 3: Empowerment of Creativity and Feedback, CD 2 were activated.

Team force. Finally, 'team force' was another element utilized for triggering CD 5: Social Influence and Relatedness, and CD 3: Empowerment of Creativity and Feedback. The biggest criticism for writing classes was their being too individual.

Students from both groups strongly complained about monotonous and individual nature of writing classes. Thus, the element of ‘team force’ was included into the design and students were given extra credits for their group studies.

Incentives and rewards. The last thing to define while designing dashboard is the rewards and incentives which would be given when students commit the desired action and/or achieved a win-state. According to Chou (2015) the values of the rewards must be in proportion to the magnitude of the win-state. In scope of the current study, the following rewards were utilized in line with the elements mentioned in previous section.

- Status points
- Badges
    - Best paper badge
    - Best feedbacker badge
    - Best task completer
    - Best starter
    - Best creator
- Positive verbal feedback from teacher
- Positive written feedback from peer
- Special consulting time

Level 1 Octalysis Ideation Process. Having completed the dashboard, the next step was to go over the eight core drives (Chou, 2015)

CD 1: Epic Meaning and Calling. One of the most difficult CDs of Octalysis framework to apply is CD 1 because it is very sophisticated and requires dedication. According to Chou (2015) one of the most critical thing to achieve an increase in CD 1 is to get participants to believe in themselves. In the case of current study, many activities and channels were embedded into the design to have students believe that they could write in the target language. Frist of all, each student had the chance to feel the accomplishment even if they had been demotivated or weak in writing competence. For this purpose, different layers which were represented by paper color codes and that were used to group students according to their achievement

levels. Thus, none of the students would feel that they had no chance of competing with their friends.

In addition, leader boards that would be indicators of students' developments were arranged in a way that each student would be evaluated with their peers at the same level. Another thing to increase the power of CD 1 was the integration of free papers. In this way, students would become more autonomous and believe in themselves more.

CD 2: Development and Accomplishment. The design of the intervention was kept simple. There were only experience points, tasks and badges. Students would not be confused by too many game elements that can be found in many games such as tens of different characters, different algorithms to calculate their experience points, many endless levels or activities in which they had to beat their opponents.

In addition, thanks to feedback mechanisms, both by the teachers and peers, which were supposed to be positive, students would never want to give up and would feel delightful each time they made a progress. The use of points and badges which were given for pre-determined achievements at all levels would also contribute to this. Students would always feel that they could have the opportunity to obtain one of the rewards once they got into action.

Another element that would have a high impact in terms of CD 2 was leader boards. As stated earlier, thanks to leaderboards, students would be aware of their developments and accomplishments. As in the use of badges of points, leaderboards arranged in a way to ensure that no matter the writing competency levels of students, they would always get the chance to improve.

CD 3: Empowerment of Creativity and Feedback. This CD was one of the most crucial ones in gamification intervention design. An increase in this CD would result in an increase in the autonomy levels of students which, in turn, contributes to CD 1 more. The use of free works which could be presented in classroom environment and could also be used to gain special consultancy time from the teachers was expected to be influential in this sense. Another important thing was that, as could be understood from the business metrics, the role of teachers was minimized in the design. Many of the works had to be carried out by students

themselves. This was important because many students (and teachers as well) were disapproving the dominant role of teachers in writing classrooms. Thus, it could be said that increased roles of students in writing processes would be beneficial for them to get their own responsibilities and become independent.

Students were also encouraged to share their works and opinions with their peers so that they would again be the ones who are responsible for their own actions and improve their self-confidences as well.

CD 4: Ownership and Possession. Offering sets of badges was an element to increase students' perception of ownership and possession. As they collect badges they would feel that they were the ones to enable that. That is, they would feel they were responsible for their own success and owned it. Moreover, integration of different kinds of badges into the design was another feature to increase sense of ownership and possession. This way, students would feel they owned their achievements at every sub-step of the process. Finally, the use of status points were another feature for CD 4. When students jumped to the next levels thanks to the status points they obtained, it would again contribute their sense of ownership.

CD 5: Social Influence and Relatedness. This was the primary objective of the current design. In the interviews teachers and students mentioned the lack of enough collaboration among students in writing classes. In addition, CD 5 was one of the weakest CDs on the Octalysis framework. So, many activities were included into the design to increase the amount and quality of the engagement between students. Experience points were given in a way that to jump to the next level a student had to cooperate with classmates, otherwise it would take a long time to collect the required sum of experience points by only individual works.

At the core of the design utilized in this study were the peer feedback mechanisms. Students were supposed to evaluate their peers' papers or works and this would enable both sides to have additional points. It was compulsory to give at least one positive feedback even the work at hand was completely written in an unsatisfyingly. In this way students not only were expected to see the positive sides but also strengthen the bounds between each other and develop positive relationships. This would also contribute the weaker students not to feel offended and in turn a decrease in peer pressure.

Moreover, task cards were including tasks which mainly required pair or group work. When students completed the tasks in pairs or groups, each student would get double experience points provided they stated their roles clearly. However, this was not limited to the task cards. Since students were free to conduct their own project in case they were not satisfied with the tasks on task cards, there was a risk of reinforcing individual studies indirectly. Therefore, team force element was included into the system so that students conducting their own projects could get the chance of winning extra points for their pair or group works.

CD 6: Scarcity and Impatience. This CD was the weakest on Octalysis framework. Therefore, it was used throughout the design process to increase the initial engagement. One of the most significant things was to motivate students who procrastinate engaging into writing activities for various reasons. Using different levels which were represented by different paper colors was one aspect of CD 6. To jump to the next level and to find out the different competition there was limited to the certain experience points. Thus, students would expected to feel the urge to collect as much points as possible to be included the higher group. Also, this was creating a challenge for students which they had to overcome to go on their journey. In this sense, levels themselves had become meaningful rewards to obtain.

In the design students were not always free. There were limitations on the number of feedbacks they could get or give. In this kind of designs, there was a risk of stronger students dominating the system. They could give all the feedbacks to the weaker students' papers. To prevent this and enable the inclusion of all students without causing any domination of a group of students, the number of feedbacks a student could give was limited to two. In the same way, it was two for the number of feedbacks to receive. When weaker stronger students filled their quotas, they had to get interaction with their weaker peers. In this way, even students did not write or want to write, they would still be interacting with their peers.

Another limitation was time. The delivery of assignments had due dates. However, these due dates did not mean that after due date the system would be blocked. Students would always have the chance to hand in their tasks, but in time deliveries were rewarded with extra experience points. In other words, due dates were used just because to enable students energized and act on time.

The utilization of several types of badges was also contributing to CD 6. They were not free and students had to overcome some challenges to get them, which would also make them valuable and scarce items. Similarly, although teachers' positive verbal feedbacks were not tangible rewards, they were valuable because to obtain them there were certain steps to complete. Only the students who brought their free tasks into the classroom could get explicit positive verbal feedback by the teacher. This was an issue which had been mentioned in the interviews by the students. some students were afraid of getting negative feedback by the teacher and as a result they did not engage in the writing activities. Although, throughout the intervention process the role of teacher was minimized, here they were encouraged to take part in processes.

Finally, another scarce and thus valuable thing was to be able to get the opportunity of special consultancy time with the teachers, which could only be accessible through the completion of certain tasks.

CD 7: Unpredictability and Curiosity. As can be seen from the Octalysis framework for writing classes of experimental group in Figure 26, CD 7 had a balanced power. It was not so weak or powerful in the classroom. This was a desired situation because excessive use of CD 7 in gamified systems could result in the emergence of undesirable consequences. On the other hand, absence of CD 7 might cause problems in terms of energizing initial actions. Therefore, although not directly utilized some use of CD 7 could be seen in the current design. First of all, task cards were offered to students in upside down way so that students could become curious about what was on a specific card and get into action. In addition, regarding the badges, students were only aware of which badges were allocated to certain tasks. For example, they were aware that if they were able to get the highest point they would get a badge from the series of 'writer's badges', or they would get one of the 'achievement badges' providing they became the best in a pre-defined category. However, they did not know which one they would get. This was aimed at increasing the curiosity power of badges.

In addition, the use of leaderboards was a source of curiosity for students because they were not aware of the sums of experience points each student got throughout the week. So, they were expected to be eager to follow boards every week.

CD 8: Loss and Avoidance. In fact, CD 8 was very powerful in experimental group and was a source of criticism by the students and teachers. Many of the students had seen it detrimental in the scope of writing classes. Therefore, in the design process, it was especially given attention to avoid the inclusion of such motivation. On the contrary, it was tried to reduce its effect. Thus, students were explicitly said that no penalties would be given whatever their desired actions would be. And in the design always alternative ways were included. For example, if students did not want to be a part of a pair or group they were free to bring their own works. Or students were even free to be inactive and were aware that they would not be faced with any kind of punishment such as low marks.

7-week intervention

The pre-intervention data which had been collected from pre-intervention survey and pre-intervention teacher and student interviews were evaluated and Octalysis frameworks illustrating the situations in both classrooms had been drawn. These frameworks gave clues in terms overall problems in classrooms. Further investigations of the data revealed that writing classes were one of the mostly criticized issues in both classrooms. Therefore, depending on the interview data, separate Octalysis frameworks were drawn and weak and strong aspects of the writing classes were identified. In line with these findings, Octalysis strategy dashboard was designed and details of the intervention process based on the eight CDs of Octalysis was determined. Finally, as shown in Figure 26, a flow chart for the intervention process was created.

Before beginning the intervention, in both groups, teachers were asked to keep logs in terms of the activities they conducted. In addition, they were asked to collect copies of students' works providing students had given consent.

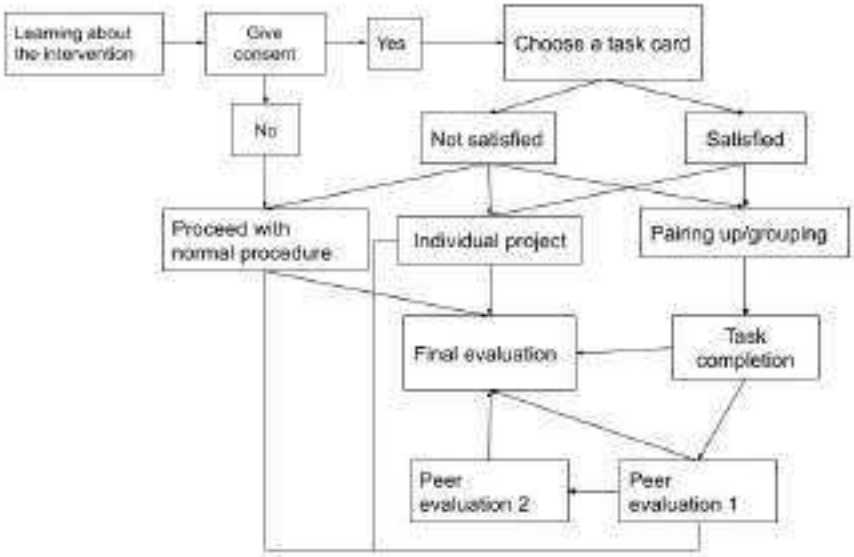


Figure 27. Intervention flow chart.

According to the flow chart, the process started with informing students in the experimental group about the intervention. And then, although they had previously given both written and oral consent for the survey, interview and intervention parts of the study separately, once more they were reminded that they were free to leave from the study and demand their data to be excluded any time they wanted throughout the intervention. All the students agreed to participate in the study. However, to ensure that participation was not compulsory, all the students were informed about the procedure in case they did not want take part in or decide to leave later. They were told that it they could go on with the routine class procedure.

Once a student gave consent, the next step was to choose one of the task cards, which were presented in an upside-down way. All the tasks cards included a specific task to write on the topic of the week, which was chosen from the coursebook that had been used at the time of study – Empower Upper Intermediate Student's Book by Cambridge University Press (2015). In accordance with the pacing of the coursebook, seven writing topics were selected. To keep the balance with the students who participated in the study and who did not or who might leave throughout the intervention, the pacing of the coursebook had not been changed. In addition, following the pacing enabled the synchronization with the control group. Table 75 shows the topics selected from the coursebook.

Table 75

Weekly Writing Tasks

	Task
Week 2	Imagine you had to live for a week without a technological device you use in your daily life. Make notes about what the experience might be like and write an article about your experience
Week 3	Choose one of the following situations and make notes on advice you could include in a leaflet. Write a leaflet for the situation you chose.
Week 4	Think about a sport or a free time activity you like or have been doing for some time. Prepare a bar chart with data on that sport or activity. Plan an activity about the data and write an article.
Week 5	Plan an email applying to do a voluntary work. Choose one of the situations given in your coursebooks (p. 53). Write an email.
Week 6	Choose one of the essay topics provided in your coursebooks (p. 65). Mae notes on possible arguments for and against. Write an essay in about 150-200 words.
Week 7	Think of a place you have visited. Plan and write a travel blog.

Having chosen a card from the deck, a student had two options. Student could go on with the task if he/she was satisfied. That is, the student was supposed to be a partner with another student or if the task had required, he could have taken part in a group. Alternatively, the student had the chance to go on with an individual project on the same topic. In case of a dissatisfaction, students had three options; going back on the routine procedure that had been carried out since the beginning of the academic year, carrying out an individual and independent project, or be a partner or a member of a group without actively taking any responsibilities. In other words, students had always had the chance to participate in the tasks any time they wanted because there might be some students who did not want to be the first to act but rather would wait and observe the process. Finally, no matter in which way students brought up their works, they were all subjected to the same evaluation. Before final evaluation students were encourage to get one or two at most peer evaluation.

Week 1. In the first week students in both groups were informed about the processes in intervention design. They were told how to take part in the activities, what were their rights, the alternative ways of completing a task, importance of working together in pairs or groups. In addition, they were also informed about the use of experience points and how the evaluation process would take place. In the last section the questions by students were answered.

Week 2. In the second week, all students in both groups had been assigned the writing task on the first topic. Students in control group did nothing different from their usual writing activities. All the students in the experimental group were writing at 'white paper level'. However, the participation from both groups was not satisfying. There were only four students who had handed in their works in the control group whereas there were six students in the experimental group.

The papers from the control group were checked by the teacher who did not prefer to give positive constructive feedback. None of the papers from the control group included any brain storming section, outline and/or plan. Moreover, there were no indication of peer feedback evidently stated on the papers, which might be one of the causes of weakness of CD 3: Empowerment of Creativity and Feedback and CD 5: Social Influence and Relatedness. The teacher of this classroom stated that this was a normal situation and found the number of students who handed in their assignments reasonable.

When the papers from the experimental group were investigated, it was seen that two out of six students wrote about their brain storming. And only one of the six students provided an outline. When the papers were examined in terms of peer feedback, there was only one student whose paper was given feedback by only one student. This could indicate that the designed system had not been able to started, yet. However, since this was the first week of the intervention, it might be too early to reach a conclusion.

Table 76

2nd Week Evaluation Results of Experimental Group

Level	ID	Brain storm	Plan	Essay	Peer Feedback 1	Peer Feedback 2	Badge
White	Student 1	-	-	50	-	-	
	Student 2	-	-	50	-	-	
	Student 3	10	-	50	50	-	Writer/Best
	Student 4	-	10	50	-	-	
	Student 5	10	-	50	-	-	Feedbacker
	Student 6	-	-	50	-	-	

As can be seen in Table 76, the number of students who actively took part in was not at desired level. That is, engagement rates were not at desired level. Besides, the interaction between students was also very limited. The problems of

CD 5: Socail Influence and Relatedness could still be observed. An answer for this came from the teacher of the classroom. According to the teacher logs, one reason for this could be the novelty of the implementation. However, the students who took part in gave positive feedback for the intervention.

On papers, it was possible to see teacher gave positive written feedback for all completed papers by using phrases such as ‘good job’, ‘good thoughts’, ‘creative’, ‘perfect’, ‘very impressing’, and so on for holistic evaluation. In addition, papers were full of grammatical corrective feedback by the teacher. Although, teacher’s positive feedback could increase the power of CD 3, focusing on grammar while giving feedback might cause some students try to avoid criticism and become demotivated or choose not to participate in the activities.

Extract 125.

Extract 125 showed that teacher still dominated the feedback processes in writing classes. The focus on students’ grammatical mistakes had heavily been criticized by the students in interviews and might have caused an increase in avoidance motivation in the classroom, which contributed to CD 8 too much.

In the 1st week, Student 3 handed in the best assignment. The paper was not only written well but also this student was more successful in terms of task completion. Therefore, Student 3 got both the writer’s badge and best paper badge. Student 5 who had given feedback for Student 3’s paper got the best feedbacker badge. Finally, Student 4 was awarded a badge for best paper. In addition, since Student 3 was the only student who accumulated more than 100 xp points. That student was upgraded to green paper level. Finally, the leaderboard was hanged on the classroom notice board.

Week 3. With the onset of task 2, the engagement levels of the students in experimental group significantly increased, which could be as a sign of success for the intervention. Because one of the primary objectives of the gamification intervention was to increase the low engagement rates in writing classes. 12 students handed in their tasks in time. Moreover, one of these students gave a second writing which was a free individual project and shared it with the rest of the classroom. In other words, the number of works delivered in week 2 had doubled. This change did not escape from the teacher's view who stated that the system could have worked.

Table 77

3rd Week Evaluation Results of Experimental Group

Level	ID	Brain storm	Plan	Essay	Peer Feedback 1 (given)	Peer Feedback 2 (given)	Badge
White	Student 1	-	10	50	50	50	Feedbacker
	Student 2	-	-	50	50	-	Creator
	Student 4	10	10	50	50	50	Best
	Student 5	10	-	50	-	-	-
	Student 6	-	-	50	50	-	-
	Student 7	10	-	50	50	-	-
	Student 8 (A)	10	10	50	-	-	-
	Student 8 (B)	10	10	50	-	-	-
	Student 9	-	-	-	50	50	-
	Student 10	-	10	50	50	50	-
	Student 11	10	10	50	-	-	-
	Student 12	10	10	50	-	-	-
Green	Student 3	10	10	50	50	-	Best/Writer

Table 77 shows the xp points distribution among the students who completed the task. As can be understood, the increase was not only in the number of students who actively took part in the tasks but also in the number of other criteria. 7 papers out of 12 had a suitable brain storming part. Similarly, seven students had planned appropriately before writing their tasks. Ten out of eleven students had written essays in line with the requirements identified in the coursebook. What was interesting was that one student (Student 9) preferred to have neither a plan nor an essay. However, Student 9 participated into the task with two feedbacks.

There were three students who did not give any feedback to the other papers whereas four students gave only one peer feedback. As a result, only four students

gave two peer feedbacks. Indeed, this could be one of the most significant improvements in writing classes of this group. Because, according to pre-intervention interviews students of this group thought that it was not easy to study with their friends in writing classed and disapproved the existing routine of writing classes. This finding was also supported by the teacher. Correspondingly, the teacher realized to eagerness of the students to write something and astounded by the high quantity of interaction among students. The teacher admitted that this was a quite uncommon inclination in writing classes. It was also acknowledged by the teacher that it had been possible to get students wrote in target language, it would have been almost impossible to create such a collaborative environment previously. In this respect, it would not wrong to state that the intervention worked. In that there was increase in CD 5: Social Influence and Relatedness, which, in turn, would contributed to an increase in CD 1: Epic Meaning and Winning.

Another interesting thing in week 2 was the two papers by Student 8. This student did not among the students who had participated into the week 1 activities. However, Student 8 brought a free individual project into classroom and shared the experience with the other students, which was a clear sign of powerful CD 3 and CD 1. For this reason, Student 8 informed that any time he/she demanded, a special consultancy hour with the teacher would be scheduled, which was an integration of CD 7: Scarcity and Impatience, and which was supposed to improve Student 8's motivation. The other work of Student 8 was partially complete writing. The brain storming and outline formation sections together with essay writing were complete. However, this student did not prefer to give or get feedback from other students.

When teacher's feedbacks were examined, it was still grammar based and focused on minor points. However, as in the first week there were positive adjectives written on students' papers as feedback. Again, it could be said that although the teacher was supporting CD 3 by using positive phrases as feedback, the focus on grammar and errors was adding value to CD 8.

As a result of the evaluation, the total scores were announced on the leaderboard. Student 4 were awarded Best paper in white category with 170 xp points. Although Student 2 did not hand in a complete task, the layout of the paper was so creative that Student 2 got the 'best creator' badge. Among the four students who had given double feedback to their friends' papers, Student 1 gave the best

feedback. Thus, ‘best feedbacker’ badge was bestowed to Student 1. As to the green level, the only student at this level was Student 3. Therefore, ‘best paper’ and ‘writer’s badges’ were given to Student 3. Finally, according to the cumulative sum of scores, Student 1, 3 and 4 got the right to jump to blue level since they collected over 200 xp points whereas Students 2, 5, 6, 7, 8, 9, and 10 jumped to the green level. This way it was possible to support CD 2 and CD 4 by using CD 7.

In terms of the works completed by the students in control group, there was not much thing to say since there were no assignments that had been delivered in the third week of the intervention. In terms of Octalysis framework, there were not considerable change.

Week 4. The upward inclination of engagement rates continued in the third week. Totally, 17 students got into interaction at varying degrees in experimental group. When the fact that the total number of students in that classroom was 20, this engagement was very meaningful. Thus, the improvement on CD 5 and CD 3 was significant. In addition, use of CD 7 by means of badges and leaderboards could have contributed CD 2, CD 4 positively. Eventually, these improvements could contribute the development of CD 1. On the other hand, there were only 7 papers which were handed in in the control group.

Table 78

4th week Evaluation results of experimental group

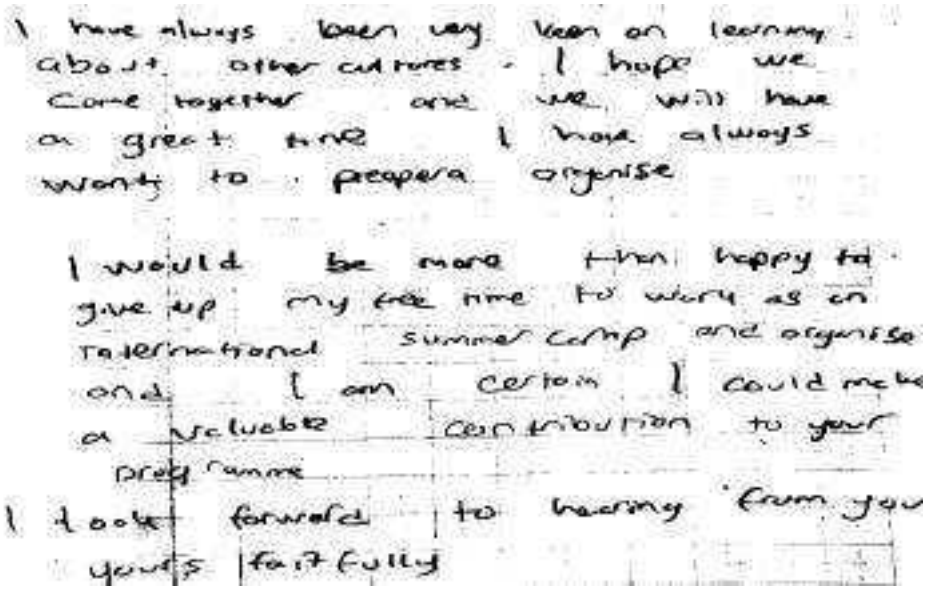
Level	ID	Brain storm	Plan	Essay	Peer Feedback 1 (given)	Peer Feedback 2 (given)	Badge
	Student 3	-	40	50	50	-	Best
Blue	Student 4	10	10	50	50	50	Completer
	Student 1		60	50	50	50	Writer
	Student 2	10	60	-	50		
	Student 5	-	10	50	-	-	
	Student 6	-	-	50	-	-	
Green	Student 7	-	40	50	-	-	Best
	Student 8	10	10	50	50	-	Completer
	Student 9		40	50	-	-	
	Student 10	10	-	50	-	-	Writer
	Student 12	-	40	50	-	-	Best
	Student 13	-	-	-	50	-	Feedbacker
White	Student 14	-	-	-	50	-	
	Student 15	-	-	-	50	-	
	Student 16	-	-	-	50	-	Feedbacker

Student 17	-	-	-	50	-	
Student 11	-	40	-	50	-	Starter

In control group, the problems with the feedback mechanisms were still evident in students' papers. One interpretation of this could be either the teacher in this group gave only oral feedback or no feedback at all.

Extract 2 presents a sample paper from control group. When it is examined, it is seen that the paper did not have any kind of feedback neither from the teacher nor from other students. That meant CD 5 and CD 3 were very low in the control group. In addition, it could be said that it was far from being satisfying, which was a sign of weak CD 2. It was a writing assignment from an upper intermediate coursebook. Apart from grammatical mistakes, major problems in terms of vocabulary selection, page layout, paragraph formation, coherence and cohesion also existed in the text. But, maybe the biggest problem of the sample was that there were no signs of interaction with peers.

Extract 126


I have always been very keen on learning about other cultures. I hope we come together and we will have a great time. I have always wanty to prepera organise

I would be more than happy to give up my free time to work as an international summer camp and organise and I am certain I could make a valuable contribution to your programme.

I look forward to hearing from you

yours faithfully

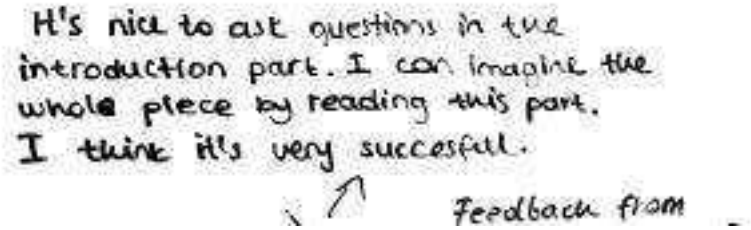
On the other hand, in the experimental group many of the students participated into the writing activities. When Table 78 is examined, it is seen that 17 students were active during the week. More complete works were presented by the students who had already been writing for the last two weeks and who were at blue or green level.

Another interesting thing was that, in week 4 five new students who had never written anything previously participated. They did not hand in a complete essay.

However, they all gave feedback to their friends. When the teacher was informed about this finding, he/she stated that those five students had never been a part of any writing activities or written anything up to that time. In this respect, to be able to integrate such demotivated students into the activities could be an important achievement of the designed gamification system. Moreover, it was evident that there was an increase in CDs 2, 3, 4, and 5 which would contribute the development of CD 1.

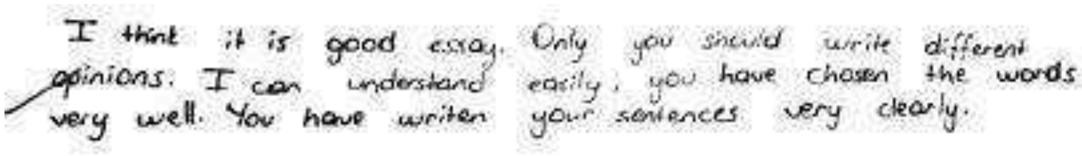
Another important thing that could be seen in Table 78 is that, students were carrying out paired and/or group activities more than the previous weeks. Because the points they got could only be gained through participating into such kind of activities. Otherwise, students could only get 10 or 50 for routine task completion. So, they had the chance to earn more xp points by getting into interaction with their friends. That is the use of points were meaningful for students and created a value. This strengthened the power of CD 2, 5 and 4. For example, In the case of Student 9, a group of six students contributed to the final work. In accordance to the task card, the outline of the work was written by Student 11 whereas the introduction part was written by Student 7, and conclusion part was written by Student 6. Then, Student 9 wrote the body paragraphs in line with introduction and conclusion and completed the final essay. Finally, Student 1 evaluated the paper and wrote a feedback.

Extract 127.



Extract 127 shows a part of feedback written by Student 1, which was quite constructive and positive. Student 1 not only indicated the level of achievement but also emphasized the stronger side of the group work. It was possible to find similar group works which had been carried out by several students and which included constructive peer feedback.

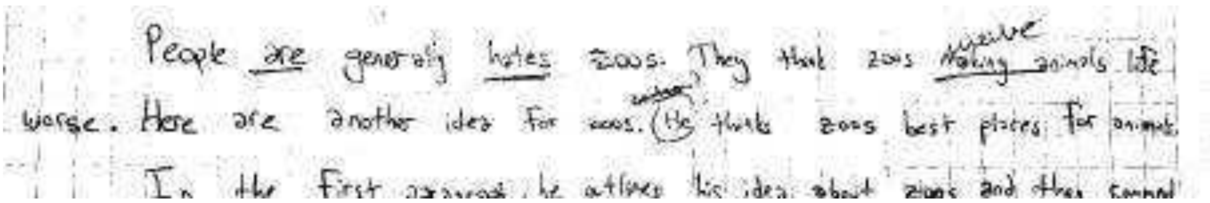
Extract 128.


I think it is good essay. Only you should write different opinions. I can understand easily, you have chosen the words very well. You have writen your sentences very clearly.

A similar type of peer feedback could be seen in Extract 128 which shows that students gave constructive feedback. The student who gave above-mentioned feedback stated his/her thoughts in a very naive way and instead of focusing on the negative parts and grammatically problematic issues, the student suggested a solution to improve the work. This might imply that students did not want to improve avoidance motivation (CD 8) and wanted their peers to strive for achieving their goals.

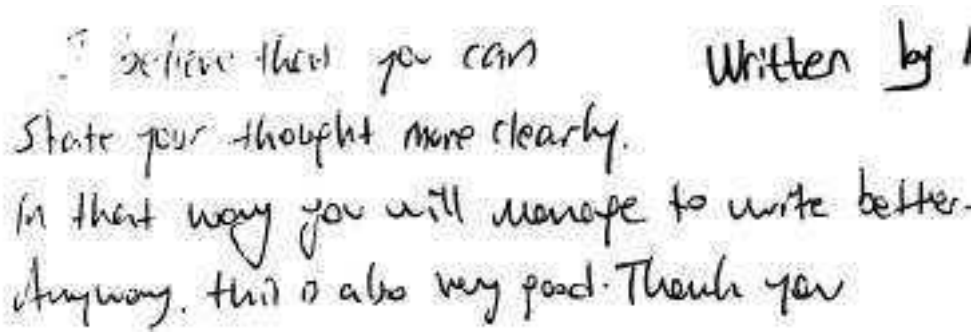
Another finding of this week was that, there was a decrease in teacher's routine feedbacks. Although the teacher evaluated some papers in terms of grammatical mistakes, in week four it was possible to see a different approach by the teacher. In many papers, the teacher preferred to give holistic feedback, which could sometimes harsh and sometimes very positive. The decrease in teachers' dominant role in the classroom was important because it put the emphasis on CD 1 and decreased the power of CD 8: Loss and Avoidance. That meant less negative feedback that focused on grammar was needed.

Extract 129.


People are generally hates zoos. They think zoos making animals life worse. Here are another idea for zoos. He thinks zoos best places for animals. In the first paragraph, he outlines his idea about zoos and then

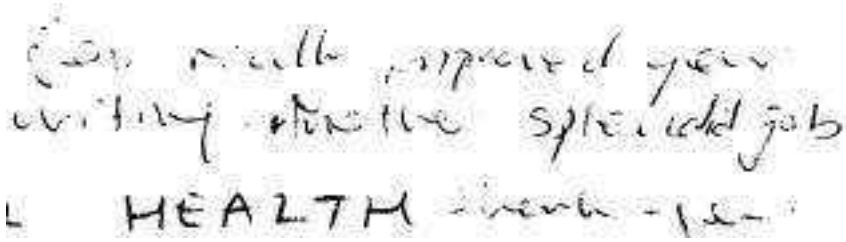
In Extract 129, it could be seen that teacher still corrected grammatical mistakes of the student. But it was evident that the teacher had decreased the number of corrections. Different from the previous weeks, this time it could be seen that teacher skipped some grammar mistakes such as the one in line 2 of Extract 5. Although students used 'are' instead of 'is', this wasn't corrected by the teacher. Similarly, in the same line the student used the superlative form in a wrong way and did not included verb to be. Again, these were not corrected by the teacher, which showed that the teacher embraced a less central role but rather a guidance role.

Extract 130.



Another example of a change in teacher's attitudes could be seen in Extract 130 which was taken from a paper on which no grammar correction existed. A holistic feedback was given by the teacher. However, different from students' modest feedbacks, teacher gave rather negative one. The use of phrases 'more clearly' or 'better' signaled that the paper was not written clearly enough and not so good.

Extract 131.



But not all holistic feedbacks were negative. As in Extract 131, there were positive feedbacks as well. The reason for teacher's adopting a different approach might be due to the increased amount of group work. Since students started to work on the projects together, it was possible they got immediate corrective feedback from their peers, and this might decrease need for teacher's feedback. That is, development in CD 1, 2, 3 and 5 was in the desired direction. However, teacher's continuous use of rather negative language might be seen as a problem since it might increase CD 8.

At the end of the week, the badges were given to the students for their outstanding achievements at each level. A total of ten badges were given. Two of them were writer's badges. At white level, there was not any satisfying task completion, so this badge was not given there. In addition, two 'best paper' badges were given, one in blue and one in white category. The use of CD 2 and 4 through CD 6 and 7 was producing satisfying results.

Week 5. In week five, there were similar situations in both groups. The number of students in control group who completed the writing assignment was five. And these were the same students who had participated into the writing activities previously. That is, there were no new students, and the majority of the classroom was inactive and preferred not to take part in writing activities. In this respect, it could be stated that the writing activities assigned by the teacher were not helping the development of CD 2, 3, 4, and 5. As a result, it was not possible to mention any improvements in CD 1 expect for the five students who were active. And there were no signs of feedback mechanisms by the teacher. Consequently, no change was expected on the Octalysis framework of this group.

When the situation was examined in experimental group, it was seen that there was a little decrease in the engagement rates. As shown in Table 79, ten students participated into the activities. But it should be noted that two students had been dropped out from the program due to their attendance problems in week 5. Therefore, the number of actively participating students was more than the half of the students in that classroom.

Table 79

5th Week Evaluation Results of Experimental Group

Level	ID	Brain storm	Plan	Essay	Peer Feedback 1 (given)	Peer Feedback 2 (given)	Badge
Yellow	Student 1	-	10	50	50	50	Best
	Student 4	10	10	50	50	50	Writer
Blue	Student 2	-	-	50	50	50	
	Student 3	-	10	50	50	50	Completer
	Student 6	-	-	50	50	-	Feedbacker
	Student 8	10	10	50	50	-	Starter
	Student 10	10	10	50	50	50	Writer
Green	Student 5	-	-	50	50	50	Feedbacker
	Student 11	-	-	-	50	-	
White	Student 17	10	10	50	50	-	Writer

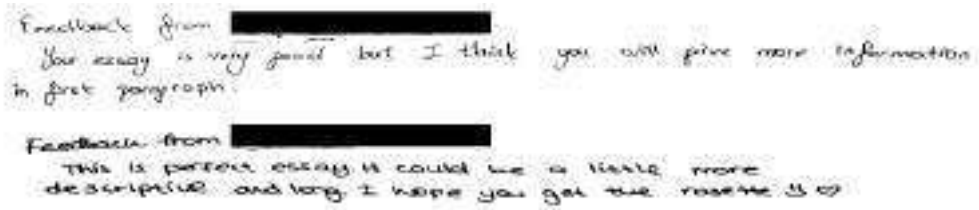
As seen in Table 79, students at blue and green levels were the ones who continued to take part in pair or group works and completed their tasks in a way. Similarly, students at white level were also attended writing activities in the previous week. In this respect, it could be stated that the gamification intervention was successful in terms of creating a sustainable environment in which CD 2, 3, 4, 5 and 6 worked well enough. In theory, it was assumed that once students felt they were

developing when they worked with their friends together and started to take their own learning responsibilities. In addition, it was possible to utilize the power of CD 6, which would contribute the development of those CDs and eventually CD 1. That is, the design of gamification intervention working in the desired way.

Eight out of ten students were awarded a badge. According to the teacher who kept logs throughout the process, this was a very popular reason for students to write. There were students who actually competed just to complete their collection of badges. And the teacher also added that it was possible to see badges attached onto the backpacks of students. Moreover, it was seen that students were using them as symbols of status in the classroom environment and seemed to enjoy the writing classes more than ever before. This finding showed that the use of points, badges and leaderboards were effective to improve the long-term motivation of students. Trying to complete the collection of badges and starting to use them as a symbol of status could be a sign of strong CD 4, 5 and 6.

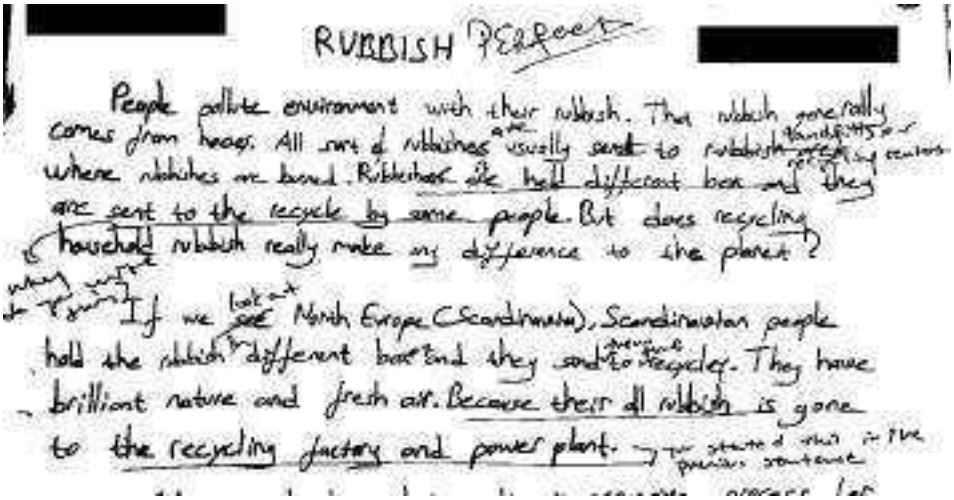
Similar issues could be observed when the content of the papers were investigated. Feedback mechanics by students and teacher were used as in the previous week. In this sense, there was not a big change. Extract 8 shows two peer feedbacks for a paper. Although both students who had given feedback thought that the paper needed some improvements, they first praised the effort of the writer with positive words and indicated what kind of improvements could be carried out. In addition, the second feedback showed how important the badges had become among the students. Indeed, the words ‘rosette’ which is frequently used as ‘rozet’ in Turkish referred the word ‘badge’. And it seems that it was important to gain one of the badges.

Extract 132.



The teacher gave feedback in a similar way. There was over-emphasis on grammar mistakes for some of the papers.

Extract 133.



RUBBISH Perfect

People pollute environment with their rubbish. That rubbish generally comes from house. All sort of rubbishes are usually sent to rubbish [illegible] where rubbishes are buried. Rubbishes like hold different box and they are sent to the recycle by some people. But does recycling household rubbish really make my difference to the planet?

If we look at North Europe (Scandinavia), Scandinavian people hold the rubbish different box and they send to recycle. They have brilliant nature and fresh air. Because their all rubbish is gone to the recycling factory and power plant.

Extract 133 shows a sample teacher feedback in the experimental group. Although the teacher used the word 'perfect' to define the quality of the study, there were many corrective feedbacks for grammar errors, vocabulary selection, or syntactic problems. The original work from which Extract 8 was taken was two-page length and full of underlined sentences that indicate a problem and/or corrected phases that showed the correct usage. In this respect, it might not be possible to mention the effect of positive feedback. The over-use of corrective feedback could result in an increase in CD 8 and create problems in terms of sustainable motivation.

Week 6. In week six, there were 12 students out of 18 who participated in writing activities in the experimental group whereas there were only two students who participated in control group. That is, the situation was not so different in the experimental group. Many of the students interacted with each other and this might indicate the continual effectiveness of the design. As students felt that they were getting more competent each week and improving their writing skills, they became more motivated and continued to take part in writing activities. In addition, high number of interactions could be seen as a sign of powerful CD 5 and 3 in the classroom.

On the other hand, there was a decrease in the number of completed tasks in the control group. According to the teacher of this classroom, the reason for this could be the approaching end-of-year proficiency exam. The teacher of the control group stated that this was a normal tendency which happened every year. As the end of year got closer, students started to focus on proficiency exam and lost their

motivation in writing. This view could indicate the existence of powerful avoidance motivation (CD 8). That is, at that time it was not important to get competent in writing and learn something for students. What was important was not to get low marks from the proficiency exam and try to avoid any potential failures.

Table 80

6th Week Evaluation Results of Experimental Group

Level	ID	Brain storm	Plan	Essay	Peer Feedback 1 (given)	Peer Feedback 2 (given)	Badge
Yellow	Student 1	10	10	50	50	50	Writer
	Student 2	-	-	50	50	-	Feedbacker
	Student 3	10	10	50	-	-	Starter
	Student 4	10	10	50	50	50	Best
	Student 5	10	10	50	-	-	Completer
Blue	Student 6	-	-	50	-	-	
	Student 7	-	10	50	50	-	Best
	Student 8	-	-	50	50	-	Feedbacker
	Student 10	10	10	50	50	50	Writer
	Student 11	10	10	50	-	-	Starter
Green	Student 9	-	-	50	-	-	Writer
White	Student 15	-	-	-	50	-	Starter

When Table 80 is examined, it can be seen that students went on writing in their desired ways. Student 6 at the blue level and Student 9 at green level preferred to write only the essay and conduct individual studies. On the other hand, Student 15 just gave a feedback to one of classmates. This was an important performance for Student 15 as this student could hardly take part in writing activities. According to the teacher, Student 15 was relatively weaker than his/her classmates. Therefore, even such small steps could mean a lot for the student. Other students were more active. They engaged in pair or group activities and interacted with each other more. In this respect, it could be argued that the gamification intervention based on Octalysis framework was a useful way to sustain high engagement rates in writing classrooms.

Week 7. As the end of term approached, the motivation in the control group decreased sharply. There were no papers handed in. The teacher of the control group stated that towards the end of the academic year this was a very normal issue in writing classes. Generally, students had already started to study for the exam and

they got highly demotivated in writing. On the other hand, there were 14 students who actively engaged in the writing tasks at the same week.

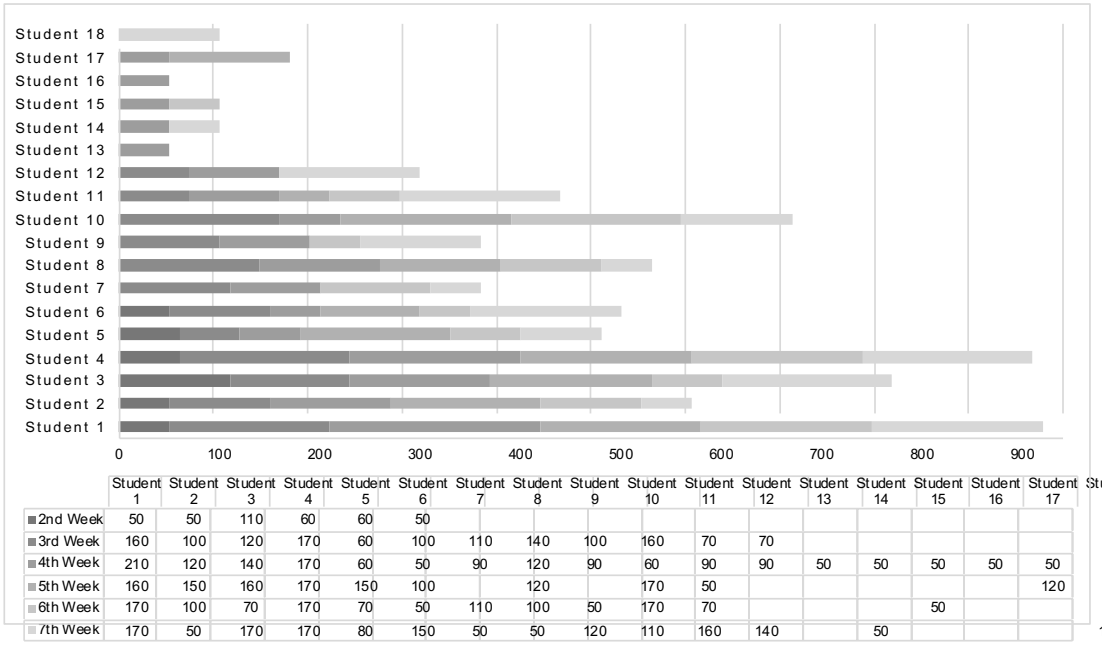
Table 81

7th Week Evaluation Results of Experimental Group

Level	ID	Brain storm	Plan	Essay	Peer Feedback 1 (given)	Peer Feedback 2 (given)	Badge
Yellow	Student 1	10	10	50	50	50	Writer
	Student 2	-	-	-	50	-	
	Student 3	10	10	50	50	50	Best
	Student 4	10	10	50	50	50	
	Student 5	-	-	80	-	-	Starter
	Student 8	-	-	-	50	-	
	Student 10	-	-	60	50	-	Feedbacker
Blue	Student 6	-	-	50	50	50	Feedbacker
	Student 7	-	-	50	-	-	
	Student 11	-	10	50	50	50	Writer
Green	Student 9	10	10	50	50	-	Best
	Student 12	10	-	80	50	-	Writer
White	Student 18	-	-	50	50	-	Feedbacker
	Student 14	-	-	-	50	-	Starter

Half of the students who took part in activities this week were at yellow level. It could be seen that at yellow level, students were able to complete their tasks better. Three of them handed in whole tasks. Similarly, four students were at the blue level, and they also managed to develop their writing skills. Moreover, since each level was evaluated separately, ten out of 14 students were able to obtain a kind of badge and found a place on the leaderboards. Even two students, Student 1 and Student 4, managed to pass the 900 cut off point and jumped to the red level.

The high interaction level among students indicated that the intervention was successful in creating a social environment. This could be seen when the overall performances of the students were investigated. Distribution of experience points and changes in students' levels illustrated how students developed their writing skills in terms of the experience points they gained.



	Student 1	Student 2	Student 3	Student 4	Student 5	Student 6	Student 7	Student 8	Student 9	Student 10	Student 11	Student 12	Student 13	Student 14	Student 15	Student 16	Student 17	Student 18
2nd Week	50	50	110	60	60	50												
3rd Week	160	100	120	170	60	100	110	140	100	160	70	70						
4th Week	210	120	140	170	60	50	90	120	90	60	90	90	50	50	50	50	50	
5th Week	160	150	160	170	150	100		120		170	50						120	
6th Week	170	100	70	170	70	50	110	100	50	170	70				50			
7th Week	170	50	170	170	80	150	50	50	120	110	160	140		50				100

Figure 28. Distribution of experience points and students' weekly engagements

As can be seen in Figure 28, it could be argued that the gamification intervention was found to be effective in terms of students' engagements in writing activities with relatively long-term motivation. Engagement levels for students in experimental group were considerably higher than the control group. When the engagement rates in the six-week period in which students were supposed carry out writing activities were taken into consideration, the average engagement rate for the experimental group was 11.83 whereas it was only 3 for control group. That is, in experimental group more than 61% of the students took part in writing activities while less than 30% of the students were active in control group.

Post-Intervention Teacher Interviews, Writing Evaluation

The data from the intervention period presented encouraging findings in terms of the effectiveness of gamification intervention in writing classes. However, quantitative data and examination of students' writings might not be enough to make conclusions. Therefore, to reach more sound conclusions, post-intervention interviews were carried out with the participation of two teachers and 10 students (5 from experimental group and 5 from control group). After seven-week intervention, the participants were asked to make similar evaluations of their writing classes as they had done previously and grade the power of eight core drives. Their comments and evaluations were investigated in accordance with the eight core drives of Octalysis.

Before the intervention, the teachers' views about the writing classes process were quite similar. According to the teachers of both classrooms, the most powerful CD affecting the motivation of their students was CD 8: Loss and Avoidance. However, interviews with the teachers after the gamification intervention showed that there were significant differences in post-intervention period.

CD 1: Epic Meaning and Calling. When the interview data was investigated, it was seen that the power of CD 1 did not change in the control group. The teacher of this group still believed that the students had moderate motivation for achieving their goals.

> Extract 134. I don't think that they changed their goals. Even, I can say that they did less as the time passed to achieve their goals. Because for the last two weeks I could not get any papers (T3).

As it is understood from the teacher's views, it was possible to mention a decrease in terms of CD 1. But the teacher of the control group decided to grade this CD as in the pre-intervention period. On the other hand, the teacher of the experimental group clearly stated that after the intervention, some of the students had shifted their goals and started to work for bigger aims.

> Extract 135. I can easily say that, students now have more meaningful goals. One of my students even attended a national writing competition and won. She now wants to attend international competitions (T1).

In the experimental group, the positive change in students' goals might be ultimate effect of gamification intervention based on Octalysis. The teacher of this group graded CD 1 higher than the previous evaluation.

CD 2: Development and Accomplishment. A similar picture could be seen in CD 2. There was not much change in the writing classes of the control group. However, the teacher of the experimental group stated that the use of points, badges and leaderboards considerably affected the motivation of students.

> Extract 136. I haven't seen such thing before. I couldn't imagine how effective these things could be. As my students collected more badges, ... they saw that they were improving. Another thing was that peer feedbacks ... really worked. They liked it (T1).

As the teacher of the experimental group indicated, the use of points, badges and leaderboards helped students see how much they improved. It seemed that these items turned into the tangible representation of their developed competences. However, previous interview data obtained in pre-intervention period had shown that

the teacher of the control group made use of the badges in the control group. But the teacher of the control group did not mention positive impact of badges after the intervention and the grade given by the teacher was not changed. That is, using badges worked in the experimental group but not in the control group. Based on this finding, it could be concluded that it was not enough just to make use of badges in the writing classes. What was important was to utilize them in a systematic way supported by the other CDs.

CD 3: Empowerment of Creativity and Feedback. When it comes to CD 3, the teacher of the control group negatively commented on the autonomous nature writing classes in which students could enjoy their own learning experiences for the sake of learning something, and took their own responsibility.

> Extract 137. I don't think there was a sustainable motivation. They gave up writing (T3).

The teacher of the control group was aware of the fact that the students did not enjoy writing activities and eventually lost their interest. On the other hand, the teacher of the experimental group pointed out that how the atmosphere in writing classes had changed.

> Extract 138. They really started to enjoy. I didn't think they would but yes writing classes were funnier for them. They waited for the announcement dates enthusiastically (T1).

The teacher of the experimental group emphasized the importance of increasing students' levels of enjoyment. Students were having fun and in turn they were participating into the activities more than ever. Another issue related to CD 3 was that the feedback mechanisms used in this group made a difference. Previously, the teacher of the experimental group stated that students' getting peer feedback caused their writing competences to improve. Therefore, it can be said that there was a difference between the two groups in terms of the impact of CD 3 after the implementation.

CD 4: Ownership and Possession. This CD was one of the CDS in which there was a change between the two groups. Before the intervention, CD 4 was very weak in the control group, and the teacher of this group stated that there was no change in this CD due to the low level of participation in writing activities. In other words, it was not possible to make any comments regarding the power of CD 4 when nobody wrote.

However, in the experimental group, the teacher indicated a slight increase in CD 4.

> Extract 139. I think I can say yes. They felt the ownership. I know students who did not put their writings away at the end of the week. I also think that those badges created such a feeling (T1).

Here as in the previous CDs, it is possible to see how different CDs were used in chunks. The use of points and badges not only contributed to an increase in CD 2: Development and Accomplishment, but also the development of CD 4. As students became more competent throughout the process, they expanded their badge collections and they started to work more to protect their possessions.

CD 5: Social Influence and Relatedness. One of the biggest differences was in CD 5. Although in the control group, the CD 5 was moderate before the intervention, the teacher of this group decreased the grade for this CD. This might have resulted from the low engagement levels of students and from the situation observed in the last two weeks.

However, in the experimental group the teacher expressed positive views.

> Extract 140. In the past I used to carry out my writing classes with the participation of five or six students. and usually towards the end of the term it used to get more difficult to maintain students' motivation high. But this time ... the number of students, especially in the last week, was astonishing. They were even asking about my views (T1).

As can be seen from the above extract, the teacher of the experimental group found the gamification intervention highly effective in terms of increased engagement levels and interaction rates. The teacher of this group stated that normally there had to be a similar situation as in the control group. But thanks to the intervention, the number of students who actively took part in activities was higher than ever. This could be an evidence of how efficiently gamification can be used.

CD 6: Scarcity and Impatience. CD 6, things were different at the beginning. In experimental group the teacher was thinking that there was no room for CD 6 in writing classes. but post-intervention interview data showed that the teacher had changed his/her mind. According to the teacher of experimental group, the using principles of CD 6: Scarcity and Impatience could yield positive results.

> Extract 141. Now, I understand how scarcity could be used. Indeed, I think, it is very effective. Limiting time, not giving awards to all and so on really helped (T1).

Since the teacher admitted that he/she had understood how CD 6 was used, maybe his/her initial evaluation of writing classes in terms of CD 6 might have not been accurate in the pre-intervention period. Yet, after the intervention, it was possible to mention an increase in CD 6 in the experimental group. This finding also indicates how important to consider the roles of different CDs while designing a gamification system. On the other hand, the teacher of the control group avoided making comments on this CD, but maintained the initial view in the pre-intervention time.

CD 7: Unpredictability and Curiosity. There was not a big change between the two groups in terms of CD 7. Both teachers had similar points of views before the intervention and they reserved their thoughts. Although the teacher of the experimental group accepted that the use of curiosity especially in task cards had a good impact on students' motivation, he/she did not change the grade for this CD.

CD 8: Loss and Avoidance. In control group, as in the pre-intervention period, the most powerful CD was CD 8.

> Extract 142. I can say that nothing has changed so far. My students were doing the minimum requirements. Actually, in the end there were no students who did the writing. I think they feared to fail in the classroom. they were studying for the final exam. So, I gave up and helped them study for the exam (T3).

As can be understood from the above extract, there was a great washback effect of proficiency exam on students' writing motivation. Since none of the students had participated in any kind of activities, it could be said that the avoidance motivation was so high that they even lost all their interest and gave up striving for their goals.

Post-intervention Octalysis graphs for teachers' evaluations

When the teachers' evaluations of writing classes before the intervention were taken into consideration, the Octalysis frameworks for post-intervention evaluation could be illustrated as in the Figures 29 and 30.

Figure 29. Post-Intervention Octalysis graph for control group, teachers' writing evaluation.

Figure 30. Post-Intervention Octalysis graph for experimental group, teachers' writing evaluation.

As can be seen in Figures 29 and 30, according to the teachers, there was a considerable difference between the two groups after the intervention. Figure 29 which illustrates the situation in the control group indicates that the avoidance motivation was still the most powerful drive that motivates students in writing classes, which might result in short-term motivation. This finding was supported with the number of students who handed in their writing assignments in the last two weeks, which was zero. That is, too much CD 8 does not allow long-term sustainable motivation. In addition, the uneven distribution of other CDs, high in black hat CDs (CDs 6, 7 and 8) and low in white hat CDs (CDs 1, 2 and 3) could cause problems in the same direction. Finally, the low levels of CD 1, 3 and 5 could be seen as significant difficulties in front of creating long-term motivation.

On the other hand, when Figure 30 is examined, it can be seen that there was a more balanced distribution of white hat and black hat drives. Especially, the low level of avoidance motivation (CD 8: Loss and Avoidance) enabled the construction of more steady atmosphere in terms of long-term motivation and creation of bigger and meaningful goals for students. Moreover, powerful CD 5: Social Influence and Relatedness and CD 3: Empowerment of Creativity and Feedback provided a suitable environment for effective intrinsic motivation. Besides, together with powerful CD 2, 3 and 5, well-adjusted use of CD 6: Scarcity and Impatience and CD 7: Unpredictability and Curiosity seemed to have contributed to the increase in CD 1, which meant that the applied gamification intervention based on Octalysis could be used to create improved DMCs for long-term motivation

Post-Intervention Student Interviews, Writing Evaluation

Post-intervention interviews with students from the control and experimental groups yielded findings similar to the ones obtained from the teacher interviews.

CD 1: Epic Meaning and Calling. Students from the experimental group indicated positive point of views in terms of CD 1.

> Extract 143. I could not get a badge but a person wants to get one indeed. But you did not write for merely get a badge. You write to learn better (S2).

From the above extract, it can be understood that Student 2 developed an understanding of how important to have bigger goals and indicated that he/she works for it.

> Extract 144. I think it was useful to achieve my goals. Normally I wouldn't have written, but I wrote just to obtain those special consulting hours with teacher because it was a privilege. The best part was the goal. It gave me a goal and encouraged me (S3).

Another student from the experimental group stated that gamification intervention helped to set meaningful goals which trigged the urge to engage in writing activities. These findings support the increased level of CD 1 in the experimental group. The average grade for this CD in the experimental group was 4.8.

In the control group the post-intervention interview results did not produce different views for this CD.

Extract 145. I don't think that anything has changed (S6).

As can be understand from Student 6's statement and from the statements of other students in the control group, it was not possible to talk about an increase in terms of meaningful goals. The average for this group was 3.6.

CD 2: Development and Accomplishment. Regarding the effect of CD 2, students in the experimental group indicated that intervention helped them feel developed.

> Extract 146. If there hadn't been a project like this, I wouldn't have written. This helped us improve our writing a lot (S2).
>
> Extract 147. Individually, I didn't give many writing tasks. But when I consider the whole classroom, it is evident that there is a development. At the beginning two of my friends had a lot of troubles in terms of their English proficiencies. But thanks to this thing not only their writing has improved but their general English too (S4).
>
> Extract 148. This also develops us. I don't throw my old papers. I am going to go over them again and develop my writing. I feel that my writing is improving (S1).

The above extracts show that students in the experimental group felt they improved their writing competences thanks to the gamification intervention. Their average for CD 2 was 5.6.

On the other hand, students from the control group did not state considerable improvement in terms of CD2. They told that in general it was possible to mention the existence of a kind of development.

> Extract 149. Since we do carry out writing tasks assigned through the coursebook, I think they have developed considerably. but I cannot say this for all of the students (S7).

However, it was not possible to make a generalization for the control group. As can be seen in the extract of Student 7, there were some students who engaged in writing activities more than their friends in the classroom but they were not many. The average for control group was 4.4.

CD 3: Empowerment of Creativity and Feedback. It was possible to witness an increase in the experimental group whereas there was a little decrease in the control group in terms of CD 3.

> Extract 150. This type of writing was better. In classical writing, students write but they don't know why. They write because they fell they have to. But now it has become more enjoyable. I write with fun indeed (S1).

According to the students from the experimental group, gamification intervention had turned writing classes into more enjoyable ones. As can be understood from the extract of Student 1, students were aware of the fact that their source of motivation changed thanks to the intervention. They felt less obliged to write and started to produce their own pieces of writings, which could be accepted as a sign of becoming more autonomous learners.

> Extract 151. You start to enjoy your own writing processes. You could feel the achievement. I realized when I spend time and effort on studying English, it pays off. The best part is I could get feedback for at least 90% of my work (S3).

A similar comment was made by another student in the experimental group. According to Student 3, the learning process itself became the source of fun, and the feedback mechanisms integrated into the gamification design seemed to have facilitated this process. The average point for CD 3 in the experimental group was 5.4. Although an increase in the power of CD 3 was seen in the experimental group, a similar inclination was not observed in the control group.

> Extract 152. ... yes, there were really creative topics in the coursebook. But somehow, I didn't went on writing. What can I say, we have an important exam ahead of us (S8).

As Student 8 from the control group pointed out, although the coursebook included creative enough topics to attract students' attention, there was something mission that had caused demotivation in the end. In addition, the impact of washback effect on students' motivation was substantial. Indeed, this finding was not contradictory with the subsequent findings of the interview of teachers and students. In the control group, the avoidance motivation was so high that it was expected to see a low level of CD 3 in this group. The average for the control group was 3.4, which indicated a two-point decrease.

CD 4: Ownership and Possession. CD 4 was one of the CDS in which a considerable change was spotted. The logic behind CD 4 by means of badges was to be able to reward students' achievements with tangible ones so that they could feel they were increasing the number of their possessions.

> Extract 153. Badges were not very important for me but one of my friends attached the badge as her first on to her pencil case, which was funny. I think this was important because when you got a badge you could feel you achieved something (S5).

Student 5 from the experimental group indicated that although the use of badges in the writing classes had not affected his/her motivation levels, he/she witnessed how the use of badges improved the sense of ownership. Accordingly, some students attached the badges they gained onto their bags or pencil cases to show they had those badges. In turn, it was expected that those students' motivation to strive for meaningful goals would be triggered and they started to study more to protect their possessions. The average of CD 4 in the experimental group was 5.2.

On the other hand, the average for the control group was 4.8, which showed a slight decrease. This might be related to the low engagement levels in control group. There were totally seven students who took part in the activities until the last two weeks. And according to the interview findings the reason for some students who had thought that CD 4 was still powerful was related to the individual participation of the activities. the average for control group was 4.8.

CD 5: Social Influence and Relatedness. Enabling an increase on CD 5 was one of the primary goals of the current gamification intervention. Low engagements levels and interactions between students were two of the mostly criticized topics in pre-intervention interview. Therefore, during the design process while creating the Octalysis dashboard, the main objective was to integrate elements that could have positive contributions in this respect. As understood from students' post-intervention comments, the elements integrated into the system had boosted the engagement and interaction levels in experimental group.

> Extract 154. Double xp points offered by task cards helped us motivate our friends who had not written anything so far. we tried to get them write to get more points, which also help them get points. ... It is for sure that group works increased the interaction in the classroom. But I don't think that feedbacks helped me improve my writing do much. But they helped to break the monotonous atmosphere in the classroom (S4).

Student 4 from the experimental group stated that the 'experience point' system helped energizing demotivated students. More active students who had already started to participate into the activities forced their friends to be included in to their tasks to get the most experience points. In this respect, even less active students did not directly dive into the writing tasks, they started to be a part of writing processes, which could be seen as a good starting point. The same student also argued that group works in writing classes resulted in an increase in interaction

levels. Finally, even though Student 4 criticized the use of feedback mechanisms impractical, he/she was sure that feedback mechanisms certainly bring out more active classes.

> Extract 155. Task cards were better than they were expected. Because they energized the ones who had not written before. Maybe we forced those friends at first but in the end, they were happy because they were at the center of attention (S5).

Student 5 from the experimental group made similar comments and talked in favor of utilizing task cards. Accordingly, tasks given on tasks cards facilitated the inclusion of demotivated students into writing activities. Student 5 also added that in the long-run this kind of application created enjoyable environments, which in turn resulted in an increase in CD 3. The average for experimental group was 6.2.

The control group, since there was no different approach adopted by the teacher, there was not a significant change. The students from the control group had still had the opinion that writing classes were mainly consisted of individual activities and this lasted till the end of the term.

> Extract 156. we usually wrote ourselves and ask teacher check our papers. But of course, we decide to write (S8).

As can be seen, the above extract suggests that students in the control group did not have to work with their friends on a project. In addition, it is understood from the last sentence of Student 8, they did not always write.

> Extract 157. ... even the task on the coursebook requires pair work, our teacher did not make it compulsory. I don't write a lot but when I ... it is always individual (S7).

It is evident that the coursebook used in the classroom contained tasks which had to be carried out in pairs or groups. But the teacher of control group did not pay attention to such things and did not encourage the students. Thus, they always could write individually and missed the chance to learn from their peers. And since CD 8: Loss and Avoidance was highly effective in the control group, this finding was not surprising. The average of control group was 2.

CD 6: Scarcity and Impatience. It was important to use CD 6 since it was possible to cause short-term motivation easily. In the experimental group, elements such as the use of points, badges and leaderboards, time restrictions, and special consultancy hours were all related to the use of scarcity principle.

> Extract 158. Badges worked. When you get a badge it means you are successful and this becomes motivating. When I got my first badge I got very happy (S5).

Student 5 pointed out that use of badge system created more motivation and joy. Gaining a badge requires successful completion of tasks and investing effort, which meant that as the award got rare, it became more valuable. This might help to trigger Students' motivation. The average for experimental group was 4.4.

In the control group, the use of scarce items in writing classes was rather low. And as in the previous CDs, there was not much change in terms of how effective the use of CD 6: Scarcity and Impatience. Students from the control group stated that the badges distributed by the teacher for successful tasks completion was important. The average for this group was 2.4, which was quite close the pre-intervention evaluation of the same CD.

CD 7: Unpredictability and Curiosity. There were not many differences for both groups in terms of the effect of CD 7. The average for the experimental group was 4.8 whereas it was 3.6 for the control group.

> Extract 159. There were many students who did not want to write but thanks to these cards they want to write. They aroused our curiosity (S1).

Student 1 from the experimental group thought that the use of task cards was an effective way of creating an attractive environment. Students who got curious to see what was in the card went on working on the task card. This might be a signal of powerful utilization of CD 7. On the other hand, students from the control group did not comment on this CD.

CD 8: Loss and Avoidance. CD 8 was the most critical CD for the success of the gamification intervention because before starting intervention, both teachers and students in both groups indicated that CD 8 was very powerful in their classes. And it was known that high CD 8 had negative impacts on long-term sustainable motivation. This could explain why all students in the control group gave up writing in the last two weeks because as can be understood from students' comments, what was important was not to fail in the classroom.

> Extract 160. ... proficiency exam was more important. I had to study for the exam. I don't want to lose another year in preparatory class. So I prefer not to spend my time for writing activities (S9).

According to Student 9 from the control group, the aim was not to achieve meaningful goals but to be able to finish the preparatory program successfully and go on with the faculty. In other words, not to fail. This was an indication of high avoidance motivation in the control group.

But in the experimental group, the students did not mentio avoiding failure so much. Their focus was on the increased interaction, improved sense of development and ownership and how joyful the classes became. Of course, like their friends in the control group, students from the experimental group were anxious about the exams and had fear of failure. But this was not specific to writing activities. Moreover, the high levels of engagement and interaction even in the last weeks indicated that avoidance motivation was not so effective in the experimental group. This was reflected in students' grading of CD 8. The average for the control group was 5.8 whereas it was 5.8 in the control group.

Post-intervention Octalysis graphs for students' evaluations

In line with the above findings, the post-intervention Octalysis frameworks for both the experimental and control groups which show students' evaluations of writing classes are as in the following.

When Figures 31 and 32 are examined, it can be seen that there is a significant difference between the views of students from both groups. In Figure 31, which illustrates the views of students about the activities carried out in writing classes, there is an uneven distribution of CDs on the octagon. The most powerful CD was CD 8: Loss and Avoidance, which is not a desired result since it results in short-term motivation. Also, vertical distribution of the CDs could be seen as problem since it means, there is little interaction between students.

Figure 31. Post-Intervention Octalysis graph for control group, students' writing evaluation.

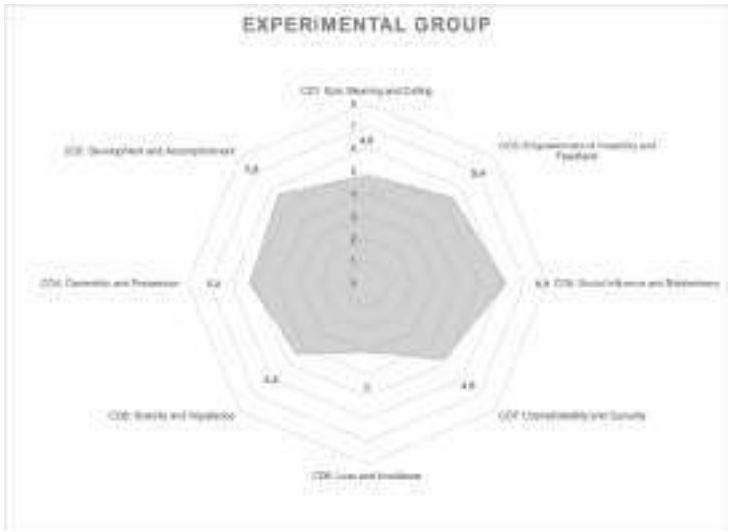


Figure 32. Post-Intervention Octalysis graph for experimental group, students' writing evaluation.

However, Figure 31 provides a more balanced distribution of the eight CDs on the octagon. The power of left hand CDs (CDs 2, 4 and 6) and the power of CDs on the left (CDs 3, 5 and 7) were almost equal, which is a desired situation in many circumstances. The most important thing in Figure 32 is the decreased level of CD 8. All in all, the improvement on these CDs is reflected by an increase in CD 1, which means that students started to study for bigger and more meaningful goals in terms of writing.

Post-intervention Overall Octalysis graphs

In the post-intervention period, the views of teachers and students have been examined by drawing visual representations of their opinions. To obtain a sounder visualization, the averages of students' and teachers' gradings which are shown in Table 81 were calculated.

Table 82

Students' and Teachers' Post-Intervention Evaluations of Writing Classes

	Experimental group			Control group		
	Students' evaluation	Teacher's evaluation	Mean	Students' evaluation	Teacher's evaluation	Mean
CD 1: Epic Meaning and Calling	4.8	5	4.9	3.6	4	3.8
CD 2: Development and Accomplishment	5.6	6	5.8	4.4	3	3.7
CD 3: Empowerment of Creativity and Feedback	5.4	5	5.2	3.4	3	3.2
CD 4: Ownership and Possession	5.2	4	4.6	4.8	2	3.4
CD 5: Social Influence and Relatedness	6.2	7	6.6	2	2	2
CD 6: Scarcity and Impatience	4.4	4	4.2	2.4	4	3.2
CD 7: Unpredictability and Curiosity	4.8	5	4.9	3.6	4	3.8
CD 8: Loss and Avoidance	3	3	3	5.8	7	6.4

When Table 82 is examined, it is seen that there is a consistency between the students' and teachers' evaluations of writing classes within groups. In the experimental group, the biggest difference was about CD 4. In this group, the students give 1.2 higher points than the teachers. However, when the values for the control group are investigated, it is seen there are significant differences between the students and teachers in terms of CD 2, CD 4, and CD 8, and the biggest difference was on CD 4.

Table 82 also provides mean scores for the experimental and control groups which are used to draw overall post-intervention Octalysis graphs for the experimental and control groups.

Figure 33. Post-Intervention Octalysis graph for experimental group, overall writing evaluation

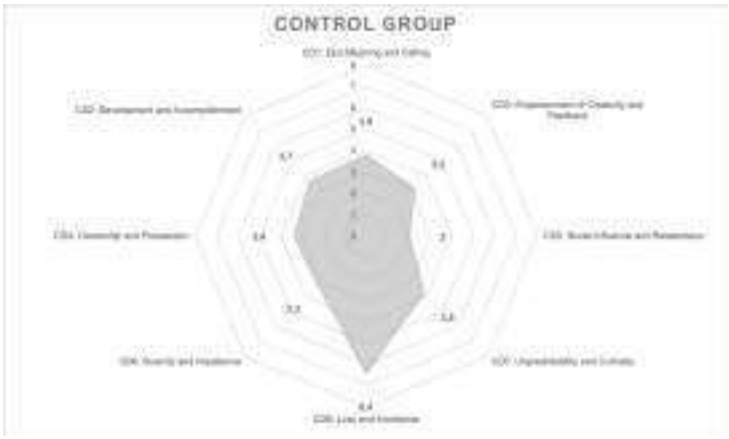


Figure 34. Post-Intervention Octalysis graph for control group, overall writing evaluation

Figure 33 and Figure 34 illustrate the post-intervention overall evaluation of writing classes. The biggest difference between the two groups was the high value on CD 8: Loss and Avoidance, which means that students in the control group are motivated just to pass their classes and not to fail in the exam. Usually this kind of motivation results in short-term motivation. The low engagement and interaction levels in the control group confirm this finding. Students who are afraid of being unsuccessful has shifted their focus on to the final proficiency exam and abandoned their writing tasks. In addition, this kind of motivation leads students to not doing their best in an activity and get satisfied with the minimum, which can be an explanation for individual and superficial task completion.

Another issue regarding the distribution of eight core drives is that in the control group, CDs are unevenly distributed on the octagon. Especially, the

unbalanced distribution of left and right CDs can be seen as the source of demotivation. Finally, the low levels of white hat core drives and uneven distribution resulted in low level of CD 1.

In the experimental group, however, a more equal spread of the CDs can be observed. As can be seen in Figure 34, the decrease in the CD 8: Loss and Avoidance is important. It allows improvement on the other CDs and eventually results in an increase in CD 1.

Chapter 5
Conclusion, Discussion and Suggestions

Introduction

In this chapter, the findings which obtained throughout the quantitative and qualitative data analyses within the extend of study are summarized and evaluated in relation to the research questions posed in the first chapter of the study. In addition, the results are investigated in terms of how they relate to the larger field of research. Moreover, the implications of the findings and the limitations of the study are examined and suggestions for further research are presented.

Discussions of the findings of the first phase

In the first phase of the study the data were collected at three stages. The first two of these (Study 1 and Study 2) included EFA procedures whereas the third study which was the main study consisted of EFA, CFA and SEM analyses.

Within the context of Study 1 (n = 115) and Study 2 (n = 338), and in the first part of the main study (n = 1670) three separate EFA tests were carried out. The goal was to investigate the validity of the selected tools and to decide whether it was possible to use them as suggested by their authors to measure the eight core drives of the Octalysis framework as it was stated in the first research question.

Discussion of the findings from EFAs. Prior to EFA tests, data sets were also examined for any violations of assumptions of normality and linearity. Thus, Skewness and Kurtosis values were calculated and investigated to ascertain that the data normally distributed. In study 1 most of the missing values were below between ±1 and some were between ±2. There were only two values out of ±2 limit. In Study 2, regarding the skewness values only four values were out of ±1 range, and there were eight Kurtosis values out of ±1. Moreover, none of the Skewness and Kurtosis values were higher than ±2. In Study 3, most of the variables had Skewness values less than ±1 and only 7 values were between ±2. In addition, all Kurtosis values in the main study were between ±2.

When the published literature was investigated, it was seen that there were several suggested cut off points to determine the normality of the data distribution.

Çokluk et al. (2012) indicate that Skewness and Kurtosis values between ±1 can be accepted. Hair et al. (2013) argue that at .05 significance level ±1.96 can be used to determine normality. There are other researchers who adopted looser limits. George and Mallery (2013) and Civelek (2017) point out that values within the range of ±2 can be used whereas Sposito et al. (1983) say it should be ±2.2. Even Kline (2011) argues that with ±10 Skewness and Kurtosis values could be the indicators of normally distributed data. So, when the findings were compared to the published limits, it was decided that the normality assumption was not violated in all Studies and it was possible to use the data sets in EFA procedures.

Besides, all data sets were examined for any nonlinear distribution of the data. For this purpose, scatterplots were used. However, due to the high number of variables, it was not possible to control all the scatterplots. So, as suggested by Tabachnick & Fidell (2013), only some scatter plots illustrating the relationships between the variables that had highest Skewness values in each data set were controlled. The findings did not provide any evidence for the linearity of the data. However, it was not possible to mention about the existence of any curvilinear distribution. Therefore, it was decided that linearity assumption was not violated in all studies. And EFA processes were conducted in the next step.

In each Study four separate EF analyses were carried out. First, scales which were adapted to measure CD 1: Epic Meaning and Calling (Muir, 2016), CD 6: Scarcity and Impatience (Kaptein et al., 2012) and CD 4: Ownership and Possession (Astaryan et al., 2013) were tested. Then, the same procedure was applied for the adapted parts of Basic Psychological Needs Satisfaction and Frustration Scale (Chen et al., 2015). The three subscales included in the EFA were Competence Satisfaction, Autonomy Satisfaction and Relatedness Satisfaction subscales, which were aimed at measuring CD 2: Development and Accomplishment, CD 3: Empowerment of Creativity and Feedback, and CD 5: Social Influence and Relatedness. Thirdly, the two sub dimensions (Stretching and Embracing) of The Curiosity and Exploration Inventory 2 (Kashdan et al., 2009) were subjected to EFA. These subdimensions aimed at measuring CD 7: Unpredictability and Curiosity dimension of Octalysis. Final EFA procedure was applied for the two dimensions (Mastery Avoidance and Performance Avoidance) of Achievement Goal

Questionnaire-Revised to measure CD 8: Loss and Avoidance (Elliot & Murayama, 2008).

In EFA analyses, first preliminary analyses were carried out. The data sets were scanned for the existence of significant correlations. For this purpose, bivariate correlation matrices were computed in SPPS and r values that were significant at .01 and .05 were sought. Because as Hair et al. (2013) argue, to factor analyze, there must be significant correlations of .3 or above (Pallant, 2011) between the variables. In addition, Bartlett's Test of Sphericity (Bartlett, 1954) was controlled for the significant relationships. Finally, for all individual EFA procedure Kaiser-Meyer-Olkin measure of sampling adequacy values were calculated. Findings of all Studies indicated that it was possible to carry out exploratory factor analyses.

Following ML extraction, as rotation method, Promax rotation which is classified under oblique rotation techniques (Hair et al., 2013) is preferred because it was thought that factors were correlated with each other (Finch, 2006). Promax rotation provided pattern and structure matrices which showed the item loadings, communalities, sum of squares, and percentage of total variances for full and reduced sets of variables in each EFA procedure. Initially, after the rotation matrices for full sets of variables were examined. Then, incase findings that indicated problems regarding the existence of low communalities, insignificant loadings, over loaded items or cross-loaded items existed, problematic items were eliminated. Each time one item was deleted and the analysis was rerun until clear sets of structures were attained. In the following paragraphs findings from each EFAs in each study were summarized together.

EFAs with DMC, Scarcity and Psychological ownership. In the first Study, it was possible to get clear pattern structures with reduced set of variables for Factor 1and Factor 2 in the fifth run. But, Factor 3 had still some issues. Factor 1 (Psychological Ownership) had three items which were POQ 1, POQ 2 and POQ 5 and their factor loadings were .918, .844 and .745 respectively. DMC 1, DMC 5, DMC 6 and DMC 12 with factor loadings .647, .592, .737 and .585 made up Factor 2. Based on these findings, it could be stated that these two factors achieved unidimensional structures. However, it was not possible to say the same thing for Scarcity scale. After the final turn, there were only two items Scarcity 2 (1.013) and Scarcity 3 (.407), which indicated that it was not possible to use this scale in

subsequent analyses. Factor 1 (eigenvalue = 1.413), Factor 2 (eigenvalue = 2.127) and Factor 3 (eigenvalue = 1.499) were explaining 55.990% of the total variance at the end of factor analysis.

In study 2, Factor 1 (DMC) had 11 items (DMC 1 to DMC 12) after the omission of DMC 4, and their factor loadings were .593, .522, .427, .623, .702, .577, .570, .527, .557, .462, and . 529 respectively. Factor 2 (Psychological Ownership) had four items POQ 1, POQ 2, POQ 3 and POQ 4, and their factor loadings were .820, .735, .631 and .768 respectively. Finally, Factor 3 (Scarcity) had four items Scarcity 1, Scarcity 2, Scarcity 3 and Scarcity 5, with loadings .329, .419, .700 and .528. As can be seen, Factor 3 had items with low loadings. However, since this is an exploratory analysis, Tabachnick and Fidell (2013) state that it is possible to interpret factor loadings over .32. Thus, all the items were retained. When the eigen values of the factors were examined, it could be seen that Factor 1 had an eigenvalue of 3.831, Factor 2 had an eigenvalue of 2.049 whereas Factor 3 had an eigenvalue of 1.110. They were all explaining 36.786% of the total variance.

In study 3, findings indicated that Factor 1 (Psychological Ownership) had four items with significant loadings: POQ 1(.925), POQ 2 (.803), POQ 3 (.721) and POQ 5 (.832). On the other hand, Factor 2 (DMC) had six items: DMC 1 (.669), DMC 2 (.537), DMC 5 (.575), DMC 6 (.697), DMC 7 (.408) and DMC 12 (.597). Finally, Factor 3 (Scarcity) had four items: Scarcity 1 (.496), Scarcity 3 (.576), Scarcity 4 (.427) and Scarcity 5 (.530). After final rotation, Factor 1 (eigenvalue = 3.219), Factor 2 (eigenvalue = 1.874) and Factor 3 (eigenvalue = .778) were explaining 41.935 of the total variance.

At the final stages of EF analyses, reliabilities and validities of the obtained structures were studied. For this purpose, first Cronbach's alpha values were calculated.

In Study 1, Cronbach's alpha values for Factor 1 (Ownership), Factor 2 (DMC) and Factor 3 (Scarcity) were .872, .731 and .606 respectively. In study 2, Factor 1 (DMC) had a reliability value of .830 while Factor 2 (Scarcity) had a reliability value of .545. The reliability value for Factor 3 (Psychological Ownership) in Study 2 was .827. Finally, the Cronbach's alpha values for Factor 1 (Psychological Ownership), Factor 2 (DMC) and Factor 3 (Scarcity) in Study 3 were .888, .746, and

.580 respectively. In their study Astaryan et al., (2013) indicate that their scale which included the Ownership dimension had a Cronbach alpha value of .89. In terms of DMC, Ghanizadeh and Jahedizadeh (2017) determined the internal consistency of DMC scale as .84, which was close to the finding of the current study. In this respect, it could be stated that the Factor 1 (Ownership) and Factor 2 (DMC) had similar internal consistency values as suggested by the other researchers. However, it might not be possible to say the same thing for Factor 3 (Scarcity) which had originally .82 Cronbach's alpha value in the original study of Borges et al., (2017). The scale was originally developed by Kaptein et al., (2012) for behavioral economics, and there were no other studies known by the researches at the time of the current study in which the Scarcity scale had been adapted for English as a foreign context. Therefore, although the internal consistency of Scarcity scale was low, it was kept for further analyses.

Hair et al. (2013) argue that reliability values need to be at or above .7 and values below this point should be reconsidered in scope of theory. Based on the findings, it can be argued that in all studies DMC and Psychological Ownership were reliable structures whereas Scarcity dimension not. Yet, for theoretical considerations, the Scarcity scale was retained. In addition, since factor loadings were high enough in all factor structures in three studies, it could be stated that convergent validity was achieved (Gaskin, 2018; Hair et al., 2013). And the low correlations between factors (.187, .214, and .300 in Study 1; .187, .214, and .300 in Study 2; and .311, .292 and . 390 in Study 3) indicated that factors were different from each other and discriminant validity was obtained (Farrell & Rudd, 2009; Gaskin, 2018).

EFAs with Basic Psychological Needs Satisfaction and Frustration Scale. The three subscales included in the EFA were Competence Satisfaction, Autonomy Satisfaction and Relatedness Satisfaction subscales, which were aimed at measuring CD 2: Development and Accomplishment, CD 3: Empowerment of Creativity and Feedback, and CD 5: Social Influence and Relatedness.

In study 1, Factor 1 (Relatedness Satisfaction) had four items with factor loadings .506 (Relatedness Satisfaction 1), .395 (Relatedness Satisfaction 2), .621 (Relatedness Satisfaction 3), and .779 (Relatedness Satisfaction 4), whereas Factor 2 (Competence Satisfaction) had only two items with factor loadings of .803

(Competence Satisfaction 2) and .724 (Competence Satisfaction 4). The factor loading for Factor 3 (Autonomy Satisfaction) were .360 (Autonomy Satisfaction 2), .524 (Autonomy Satisfaction 3) and .883 (Autonomy Satisfaction 4). After rotation all three factors were explaining 45.091 of the total variance and their eigenvalues were 2.869, .771 and .418 respectively. Although, eigenvalues seemed low, it should be kept in mind that while determining the number of factors 'a priori criterion' was applied (Hair et al., 2013) because it was in line with the theory.

In Study 2, Factor 1 (Autonomy Satisfaction) had three items. Their factor loadings were .387,.634 and .798. On the other hand, Factor 2 (Competence Satisfaction) had three items and their factor loadings were .363, .537, .812 respectively. These two factors were explaining 41.906% of the total variance and their eigenvalues were 2.065 (Factor1) and .449 (Factor 2). The problem with the explained variance with these findings were still evident after the Study 1.

In Study 3, Factor 1 (Competence Satisfaction) had four items and their factor loadings were .690 (Competence Satisfaction1), .470 (Competence Satisfaction2), .625 (Competence Satisfaction3) and .788 (Competence Satisfaction 4). Factor 2 (Autonomy Satisfaction) had factor loadings of .480 (Autonomy Satisfaction 1), .648 (Autonomy Satisfaction 2), .795 (Autonomy Satisfaction 3) and .568 (Autonomy Satisfaction 4). Finally, factor loadings for Factor 3 (Relatedness Satisfaction) were .468 (Relatedness Satisfaction 1), .491 (Relatedness Satisfaction 2), .762 (Relatedness Satisfaction 3) and .520 (Relatedness Satisfaction 4). Factor 1 (eigenvalue = 3.398), Factor 2 (eigenvalue = .968) and Factor 3 (eigenvalue = .554) were explaining 41% of the total variance.

The Cronbach's alpha values for each study were as follows. In Study 1, Cronbach's alpha was .659 for Factor 1 (Relatedness Satisfaction), .662 for Factor 2 (Competence Satisfaction) and .699 for Factor 3 (Autonomy Satisfaction). In study 2, they were .641 for Factor 1 (Autonomy Satisfaction) and .651 for Factor 2 (Competence Satisfaction). Finally, the reliability values for Study 3 were .774 for Factor 1 (Competence Satisfaction), .647 for Factor 2 (Autonomy Satisfaction) and .723 for Factor 3 (Relatedness Satisfaction). In literature, Chen et al., (2015) and Hu and Zhang (2017) indicate that all these subscales have a reliability value of or above .80. In this respect the findings in this current study were a little bit lower that findings of the original studies.

However, As can be seen the reliabilities of the structures increased with each study. Although in the final study Factor 2 still had a relatively low reliability. But it was not far from the limit. So, it was decided to proceed with the examination of validities. Since all items loaded highly in each structure, it was assumed that convergent validity was achieved in each study. The inter-factor correlations were .601, .460 and .649 in Study 1 whereas as it was .578 in Study 2. And in the main EFA of main Study they were .646, .352 and .432. Findings indicated that in Study 3 inter-factor correlations were relatively high, which might mean that they were measuring the same concepts. Although these values were relatively high, they were below the suggested level by Gaskin (2018) which was .7. In addition, the item loadings were high enough. Therefore, it was assumed that discriminant validity was also achieved and factors were retained.

Above-mentioned findings indicated that it was possible to use adapted version of the three subscales of BPNSF scale in scope of the current study.

EFAs with The Curiosity and Exploration Inventory 2. With the EF analyses which was carried out with CEI-II within the context of this study aimed at measuring CD 7: Unpredictability and Curiosity dimension of Octalysis. This scale originally consisted of two subscales: Stretching and Embracing.

In Study 1, Factor 1 had three items with .311 (Embracing 2), .1.035 (Embracing 3) and .514 (Embracing 4) whereas Factor 2 with three items had .590 (Stretching 1), .766 (Stretching 3) and .456 (Stretching 5). Factor 1 (eigenvalue = 2.138) and Factor 2 (eigenvalue = .923) were explaining 51.020 of the total variance.

In Study 2, Factor 1 (Embracing) had two items with factor loadings of .939 (Embracing 3) and .531 (Embracing 4) while Factor 2 (Stretching) had three items with loadings of .456 (Stretching 1), .320 (Stretching 2) and .248 (Stretching 3). The two factors whose eigenvalues were 1.553 and .534 respectively were explaining 41.746 of the total variance. These findings showed that it might not be possible to use Embracing dimension of CEI-II in SEM analyses.

In study 3, the item loadings of Factor 1 (Stretching) were .641 (Stretching 1), .709 (Stretching 3) and .606 (Stretching 5). On the other hand, the factor loadings for Factor 2 (Embracing) were .464 (Embracing 1), .715 (Embracing 3) and .616 (Embracing 4). Factor 1 (eigenvalue = 1.950) and Facor 2 (eigenvalue = .458) were

explaining 40.133% of the total variance. In this respect, Factor 2 seemed to contribute to the explained total variance very little. However, since three items loaded significantly enough onto this factor, and theoretically it was logical, it was decided to retain these factors.

The Cronbach's alpha values in Study 1 were.66 (Factor 1) and .92 (Factor 2) whereas they were .633 for Factor 1 and .498 for Factor 2 in Study 2; and .673 for Factor 1 and .588 for Factor 2 in Study 3. Ye et al., (2015) and Kashdan et al. (2009) indicate that the two subscales have a good internal consistency. But the findings of the current study contradict with theirs, especially the reliability of the subscale Embracing. Based on this finding it could be argued that in the final study there was a low reliability. However, when the process examined the problem might stem from Embracing dimension. Accordingly, it could be stated that it was possible to obtain a reliable structure with Stretching dimension of CEI-II whereas it was not possible to use Embracing dimension as suggested by the authors. Therefore, discriminant validity check was not carried out.

EFAs with Achievement Goal Questionnaire-Revised. To measure CD 8: Loss and Avoidance of Octalysis, the two dimensions of AGQ-R (Mastery Avoidance and Performance Avoidance) were adapted and tested.

In Study 1, after the rotation, it was seen that Mastery Avoidance 3 cross-loaded onto the two factors. Thus, it was eliminated and the model was respecified. It was seen that there was an improvement in the model. When the factor loadings were examined, it could be seen that Performance Avoidance 1 cross-loaded onto the two factors. However, it loaded on Factor 1 with .584 while it loaded on Factor 2 with .304. It was evident that this item strongly loaded on Factor 1 so it was not necessary to respecify the model. The eigenvalue of Factor 1 was 2.377 whereas it was .872 for Factor 2 after the final rotation. The two final factors were explaining 64.990 of the total variance.

In Study 2, Factor 1 (Performance Avoidance) had loadings of .622 (Performance Avoidance 1), .857 (Performance Avoidance 2) and .860 (Performance Avoidance3). On the other hand, Factor 2 (Mastery Avoidance) had loadings of .953 (Mastery Avoidance 1), .677 (Mastery Avoidance 2) and .436

(Mastery Avoidance 3). Factor 1 (eigenvalue = 2.389) and Factor 2 (eigenvalue = 1.145) were explaining 58.900% of the total variance.

In Study 3, Factor 1 (Performance Avoidance) had three items, and their factor loadings were .755 (Performance Avoidance 1), .829 (Performance Avoidance 2) and .877 (Performance Avoidance3). Factor 2 (Mastery Avoidance) had three items, too, and their factor loadings were .588 (Mastery Avoidance 1), .783 (Mastery Avoidance 2) and .489 (Mastery Avoidance 3) respectively. Factor 1 (eigenvalue = 2.516) and Factor 2 (eigenvalue = .737) were explaining 54.209% of the total variance.

Since in all items loaded onto one factor in all studies, it could be said that unidimensionality was achieved in Study 1, 2 and 3. In addition, Cronbach' alpha for Factor 1 was .804 whereas it was .801 for Factor 2 in Study 1. On the other hand, they were .833 (Performance Avoidance) and .722 (Mastery Avoidance) in Study 2. Finally, performance Avoidance Factor had .862 reliability whereas Mastery Avoidance Factor had .646 reliability. These values were partially in line with the findings of the Apostolou (2013) who stated that these two dimensions have a high reliability values of above .90. Nevertheless, these findings indicated that in all studies reliable structures were attained. In terms of convergent validity, items loaded onto factors strongly, which as an indication of convergent validity. Moreover, Inter-factor correlation was .454 in Study 1, .376 in Study 2, and .512 in Study 3, which were not indicators of perfect discriminant validity but low enough to assume that discriminant validity was achieved.

Research question 1: Can preselected scales be used as suggested by their authors to measure the eight core drives of Octalysis? Within the extend of first phase of the current study three different EFA procedures were applied and it was aimed to explore the usability of the selected scales in accordance with the aims of the current study. Table 83 summarizes the items in each factor that had significant loadings and reliability values for each factor were obtained in each study.

Table 83

Reduced Sets of Items for Octalysis in Study 1, Study 2 and Study 3

Scale	Subscale	Items in Study 1	α	Items in Study 2	α	Items in Study 3	α
DMC	-	1, 5, 6, 12	.731	1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12	.830	1, 2, 5, 6, 7, 12	.746
STPS	-	2, 3	.606	1, 2, 3, 5	.545	1, 3, 4, 5	.580
POQ	-	1, 2, 5	.872	1, 2, 3, 5	.827	1, 2, 3, 5	.888
BPNSF-General	Autonomy Satisfaction	2, 3, 4	.699	1, 2, 3	.641	1, 2, 3, 4	.774
	Relatedness Satisfaction	1, 2, 3, 4	.658	-		1, 2, 3, 4	.723
	Competence Satisfaction	2, 4	.662	1, 3, 4	.651	1, 2, 3, 4	.647
AGQ-R	Mastery Avoidance	1, 2	.801	1, 2, 3	.722	1, 2, 3	.646
	Performance Avoidance	1, 2, 3	.804	1, 2, 3	.833	1, 2, 3	.862
CEI-II	Stretching	1, 3, 5	.660	1, 2, 3	.498	1, 3, 5	.673
	Embracing	2, 4	-	3, 4	.633	1, 3, 4	.588

As can be seen in Table 83, In Study 1, after Embracing dimension of CEI-II was eliminated, it was possible to use the other scales with their reduced sets of variables. However, the subscales of Scarcity, Competence Satisfaction and Mastery Avoidance which had two items would have been underidentified if they had been used in a SEM analysis. When the number of observations which was 92 was taken into consideration, it would not be wrong to state that sample size might have had resulted in such finding (Hair et al., 2013; Tabachnick & Fidell, 2013).

On the other hand, in Study 2, after all EFA procedures had been completed, it was found out that apart from the subscale Relatedness Satisfaction of BPNSF-General, it was possible to utilize other subscales as suggested by their authors in reduced versions.

In Study 2, it was possible to obtain better results compared to Study 1 because there were more subscales that could be used in subsequent statistical analysis. According to the findings from Study 1, it was not possible to use Embracing subscales in CEI-II, Scarcity subscale from STPS, Competence Satisfaction subscale from BPNSF-General, and Mastery Avoidance subscale from AGQ-R in SEM analyses. But in Study 2, it was found out that except Embracing subscale in CEI-II and Relatedness Satisfaction in BPNSF-General, the other subscales could be used as suggested by their authors. The reason for this could be the adjustments that had been made prior to Study 2.

However, Study 2 was not free from problems. Findings indicate that some subscales failed to go above the suggested limits in terms of their validities and reliabilities. Scarcity (.545) and Stretching (.498) subscales had critically low levels of Cronbach's alpha.

At the end of Study 1 and Study 2, some adjustments and arrangements had been made and revised version of Octalysis survey was applied in Study 3. The findings indicated that the interventions and changes resulted in more consistent structures.

In conclusion, in relation to the first research question, it was not wrong to state that preselected scales can be used as suggested by their authors to measure the eight core drives of Octalysis.

Discussion of the findings from SEM. Findings from EF analyses summarized in the previous section provided a complete Octalysis survey which could be used to measure the eight core drives of the DMC embedded Octalysis theory. Therefore, in the next step it was aimed to investigate how well the variables could measure the constructs underlying Octalysis. The goal was to find whether there was a consistency between the data and the measured constructs, which was tested by CF analysis. Then, with the analysis of the structural model, it was aimed at finding answers for the hypotheses formed in accordance with the DMC embedded Octalysis theory.

Within the scope of SEM analysis, first individual constructs were defined. Totally 37 items which had been validated in EF analyses were grouped under 8 factors which were labeled in accordance with the eight CDs of Octalysis. Then, working definitions were done for each construct. Following this process, an overall measurement model was developed. Accordingly, the model had eight latent constructs which were measured by 37 indicators. This was a causal model and the causality was from the latent constructs to the measured items. In addition, items in each single construct were supposed to measure the similar concepts and empirically work together. The latent constructs were not independent from each other, and thus, it was supposed that they all correlated with each other. One latent construct which was CD 1: Epic Meaning and Calling was measured by six items whereas five constructs which were CD 2: Development and Accomplishment, CD

3: Empowerment of Creativity and Feedback, CD 4: Ownership and Possession, CD 5: Social Influence and Relatedness, and CD 6: Scarcity and Impatience were indicated by four items. CD 7: Unpredictability and Curiosity was indicated by five items while CD 8: Loss and Avoidance was indicated by six items which were grouped under two subdimensions.

After the identification of the model, it was seen that the model had 112 estimated parameters 96 of which were free parameters, 28 of which were factor loadings, 45 of which were factor variance and covariance, and 39 of which were error terms. The number of unique variance and covariance terms was 780 which was greater than 112. This finding indicated that the model was identified and had more degrees of freedom than free parameters (Tabachnick & Fidell, 2013).

In the next step, the validity of the measurement model was assessed (Ullman, 2013). To this end, the consistency between sample covariance matrix which represents theoretically developed model and population covariance matrix which represented the model in reality was controlled and Goodness-of-Fit indices were examined. As Hair et al. (2013) state provided the indices were above the suggested level, the findings could be accepted as the indicators of good fit. However, there are many goodness-of-fit indices in literature and to be able to make a decision about the validity of the model, it is necessary to control various indices (Ulmann, 2007; Bentler and Bonett, 1980; Hu and Bentler, 2013; Kline, 2011). Hair et al. (2013) suggest controlling indices from three categories which are absolute fit indices, incremental fit indices, and parsimony fit indices.

When the measurement model was examined, it was found out that the chi-square was 1243.744 with 593 degrees of freedom ($p = .000$). Unlike other statistical tests, CMIN (x2) value need to be insignificant in SEM analyses. Because significant p-values indicate that observed and estimated model are significantly different from each other. Since the higher conformity between the models means better fit, insignificant p-value is desired. In this respect, significant x^2 obtained from the measurement model could indicate a problem. However, as Hair et al. (2013) point out, when the sample size is over 250 and the number of observed parameters is more than 30, significant p-values are almost inevitable. That is, as Ullman (2013), Gulliksen and Tukey (1958), Collier (2020) and Jöreskog and Sörbom (1969) argue, there is a close relationship between the sample size and x^2 and as the sample size

increases, the probability of the insignificant x^2 decreases. When the sample size of the current study which was 1150 and the number of observed parameters 37 were taken into consideration, it could be stated that it was normal to get significant CMIN.

To overcome the problems caused by the x^2 with big sample sizes, Collier (2020) suggests using relative chi-square value (CMIN/DF). According to Carmines and McIver (1981) CMIN/DF vales between 2 to 1 or 3 to 1 ratios can indicate adequate fit. On the other hand, Byrne(1989) state that CMIN/DF value should not exceed 2 for good fit whereas Kline (2011) point out that values under 3 are the indicators of acceptable fit. CMIN/DF value for the measurement model of the current study was 2.097 which was within the range of suggested limits. Thus, it could be stated that it indicated the existence of a good fit.

In addition to the chi-square values, various indices from absolute fit measures, incremental fit indices and parsimony fit indices were also checked. The RMSEA value for the measurement model was .031 (between .028 and .033 with 90% confidence interval). According to Hair et al. (2013) RMSEA indicates fitness of the model to the population. As Kline (2011) states closer RMSEA values indicate better fit. Browne and Cudeck (1993) argue that RMSEA value below .05 indicates a close fit but values below .08 are also acceptable. .05 cutoff point is suggested by other reseachers such as Brown (2006), Jöreskog and Sörbom (1969), Sümer (2000) and Collier (2020). However, it is possible to find researchers who have adopted looser limits. For example, Thompson (2004) state that RMSEA values up to .06 can be accepted whereas Steiger (2007) the upper limit could be .07. On the other hand, Sümer (2000) and Hooper, Coughlan and Mullen (2008) values below .08 could be used as the indicators of good fit. When the finding of the current study and the suggested limits were compared, it could be said that the tested model had a good fit. In addition to the RMSEA, it is suggested to report SRMR values when the sample size is over 500 (Kline, 1998). The SRMR value for the measurement model was .036 which was below the suggested cutoff point and which indicated good fit between matrices (Hair et al., 2013).

Among the incremental fit indices, comparative fit index (CFI) which is one of the commonly used indices and which is not effected by sample size (Collier, 2020) was also examined. CFI values range between 0 and 1, and the closer the value to 1, the better fit it indicates (Hair et al., 2013). According to Bentler and Bonett (1980)

CFI values should not be below .90 for acceptable fit. And as Hu and Bentler (1999) point out for good fit, CFI should be at or above .95. The CFI value for the measurement model was found to be .949 which was above the suggested levels and indicated a good fit between the models.

To sum up, although it was possible to include many more indices, the above-mentioned indices provide sufficient evidence that support good model fit. Therefore, further analyses to determine the construct validity of the measurement model were carried out. In this respect, convergent validity and discriminant validity were inspected.

For convergent validity which shows that indicators of a specific construct measure that construct, factor loadings of the items should be controlled (Hair et al., 2013). In this sense, high factor loadings could be seen as the indicators of high convergent validity. In addition, at the minimum all factor loadings need to be significant at $p < .01$ level. When the unstandardized estimates and standardized factor loadings were investigated. It could be seen that except for the items DMC 7 (.41) and STPS 4 (.39), all the other items were ranging between .92 and .50 and all the factor loading were significant at .01 level. Hair et al. (2013) state that ideally factor loadings should be higher than .7 but loadings over .5 can also be accepted. According to Collier (2020) as a rule of thumb factor loadings need to be higher than .05. on the other hand, Garson (2010) and Stevens (1992) argue that loadings as low as .4 can be retained as weak ones. In this respect, when the complexity of the model and the theory being tested were taken into consideration, no actions were taken and all the items were retained.

Another indicator of the convergent validity is the average variance extracted (AVE) values (Fornell & Larcker, 1981). According to Collier (2020) AVE values above .5 shows that there is a convergent validity. Findings of the current study indicated that CD 2: Development and Accomplishment (AVE = .46), CD 4: Ownership and Possession (AVE = .68); and CD 8: Loss and Avoidance, Performance Avoidance (AVE = .68) had AVE values at or higher than .5, which implied the existence of convergent validity for these constructs. On the other hand, CD 1: Epic Meaning and Calling (AVE = .35), CD 3: Empowerment of Creativity and Feedback (AVE = .41), CD 5: Social Influence and Relatedness (AVE = .33), CD 6: Scarcity and Impatience (AVE = .26), CD 7: Unpredictability and Curiosity (AVE =

.40), and CD 8: Loss and Avoidance, Mastery Avoidance (AVE = .40) had AVE values below the suggested levels. However, Fornell and Larcker (1981) argue that AVE is too strict to determine the convergent validity and it is possible to use structures with AVE less than .5. Similarly, Huang et al. (2013) state that .5 for AVE is the desired level but it is possible to accept values as low as .4. Steinmetz (2016) who also criticizes using AVE values to determine the convergent validity suggests that composite reliability (CR) could be enough to make decisions in terms of convergent validity. Similarly, Kline (2011) and Malhotra and Dash (2011) state that AVE is too conservative and instead CR alone could be used. Fornell and Larcker (1981), Gaskin (2018) and Lam (2012) point out that when the AVE is less than .5 but CR is at or above .6, it is possible to argue that model has adequate convergent validity. In scope of the current study, the CR values (CD 1, .77; CD 2, .78; CD 3, .73; CD 4, .90; CD 5, .66; CD 6, .58; CD 7, .76; and CD 8, .66 and .86) obtained from the measurement model were taken into consideration, it was possible to argue that convergent validity was adequate.

All in all, it would not be wrong to state that although average variance extracted estimates seemed to cause problems in terms of the existence of convergent validity, most of the factor loadings being above suggested level and CR values being at or above .6 provided enough evidence to support the convergent validity of the Octalysis model.

To decide the validity of the model, it should also be examined in terms of discriminant validity which is calculated with the shared variance between constructs (Fornell & Larcker, 1981). The squared correlations between constructs need to be higher than the AVE of each construct. However, Collier (2020) and Henseler et al. (2015) criticize assessing discriminant validity as suggested by Fornell and Larcker (1981) by finding it unreliable, and they suggest using Heterotrait-Monotrait Ratio of Correlation (HTMT) instead of Fornell-Larcker criterion. In this respect Kline (2011) states that HTMT value above .85 is the indicator of validity problems. When the HTMT values of the measurement model were investigated, it could be seen that they all ranged between .067 and .708 and indicated the existence of discriminant validity.

Research question 2: Is it possible to measure what is intended to measure with the adapted version of the survey? When the findings from the CFA stage were

analyzed, it could be seen that the measurement model was identified and had more freedom than the free parameter. Thus, it was appropriate to proceed with further analyses. In addition, the investigation of chi-square and related data and goodness-of-fit indices for the measurement model, it was found out that the model had a good fit. In terms of the construct validity, significant factor loadings of the items and composite reliabilities of the construct provide enough evidence for convergent validity whereas HTMT ratio of correlations provide enough evidence for discriminant validity.

Thus, based on above-mentioned findings, it would not be wrong to state that the tested measurement model had construct validity and items accurately reflect theoretically constructed latent structures**.**

Structural equation modelling. The second part of SEM analysis consisted of specification of the structural model. Previously, in the measurement model the relationships between latent constructs had been investigated. In structural models, however, it was aimed to show those relationships in the theoretical model via a diagram and the magnitude of the relationships were examined. In this respect, full structural model was specified by defining the structural theory of DMC embedded Octalysis theory. Based on this theory eight hypotheses were formed as in the following: English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with:

1. their feeling of accomplishment, skills development and overcoming challenges
2. their involvement in creative processes
3. their feeling of the ownership regarding their own learning processes
4. their feeling of being related to a social group in their learning environments
5. their desire to obtain scarce things and their perception of closing deadline
6. their being curious about the things in their learning environments which get their attention and their encounter with unpredictable things
7. their endeavors not to lose something and to avoid failure

Having specified the full structural and formed the related hypotheses, in the next step the validity of the structural model was assessed. For this purpose, goodness-of-fit measures were investigated. Accordingly, x^2 was 1272.190 with 599 degrees of freedom at p = .000 level. Although significant chi-square values were the indicators of bad fit, they were known to be highly sensitive to sample sizes (Hair et al., 2013; Ullman, 2013; Gulliksen and Tukey, 1958; Collier, 2020). Thus, as suggested by Kline (2011), Collier (2020) and Carmines and McIver (1981) before giving any decisions, CMIN/DF value was controlled. It was found out to be 2.124, which was below the acceptable threshold (Kilne, 2011). The RMSEA which is another absolute fit index was .031 (90% confidence interval of .029 to .034) whereas SRMR which is an index that is suggested to be controlled by Kline (1998) when the sample size exceeds 500 and was .0382. According to Browne and Cudeck (1993), Brown (2006), Jöreskog and Sörbom (1969), Sümer (2000) and Collier (2020), RMSEA value under .05 indicates perfect fit. In addition, SRMR value under .05 signals good fit (Kline, 1998). In this respect, when these RMSEA and SRMR values were taken into consideration, it could be concluded that the model had perfect fit.

The CFI value was found to be .948. According to Bentler and Bonett (1980), Hair et al. (2013) CFI above .90 indicates acceptable model while Hu and Bentler (1999) state that for good model fit it should be above .95 level. Based on these published thresholds, it could be argued that CFI value was indicator of a good model fit. In terms of other goodness-of-fit indices, it was found out that NFI was .906, RFI was .895, AGFI was .934 and PNFI was. 815, all of which indicate good model fit. In line with these findings, it was assumed that the model had a good fit.

In the next step, standardized estimates of the loadings and path coefficients were investigated (Hair et al., 2013), and findings were compared with the ones from CFA analysis. It was observed that loading estimates did not change considerably in the SE model. These findings indicated that parameter stability was achieved which also provided further evidence for the validity of the measurement model. In addition, it was found out that consistency between the loading estimates of CFA and SEM reproduced in the consistency of construct reliabilities. In most of the constructs, CR values did not change. The only change was in CD 1 and CD 5, which was .01 drop, and which could be ignored.

To determine the validity of the model significance of the path estimates were checked. All the estimates were in the expected direction and apart from the estimate between CD 3 and CD 1, all other path estimates were significant at .05 level. Since six out of seven estimates were sufficiently significant in line with the hypotheses, it would not be wrong to assume that theoretical model was supported. Accordingly, the following decisions regarding the hypotheses were made:

H1: *English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with their feeling of accomplishment, skills development and overcoming challenges* was supported (unstandardized estimate = .161, S.E. = .069, t-value = 2.33, standardized estimate = .168)

H2: *English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with their involvement in creative processes* was not supported (unstandardized estimate = .024, S.E. = .071, t-value = .34, standardized estimate = .023).

H3: *English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with their feeling of the ownership regarding their own learning processes* was supported (unstandardized estimate = .093, S.E. = .025, t-value = 3.773, standardized estimate = .149).

H4: *English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with their feeling of being related to a social group in their learning environments* was supported (unstandardized estimate = .106, S.E. = .053, t-value = 2.003, standardized estimate = .094).

H5: *English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with their desire to obtain scarce things and their perception of closing deadline* was supported (unstandardized estimate = .156, S.E. = .048, t-value = 3.270, standardized estimate = .183).

H6: *English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with their being curious*

about the things in their learning environments which get their attention and their encounter with unpredictable things was supported (unstandardized estimate = .146, S.E. = .073, t-value = 1.983, standardized estimate = .113).

H7: *English as a foreign language students' long-term motivation to achieve big and meaningful goals is correlated with their endeavors not to lose something and to avoid failure* was supported (unstandardized estimate = .102, S.E. = .049, t-value = 2.089, standardized estimate = .091).

Above-mentioned findings provided enough psychometric evidence to assume that based on the DMC embedded Octalysis theory, it could be possible to design gamification interventions.

Discussions of the findings of the second phase

In the second phase of the study it was aimed at designing and implementing a gamification intervention based on the DMC embedded Octalysis theory. Throughout this phase quantitative and qualitative data collected together.

Quantitative findings. First, the Octalysis survey which was validated in the first phase of the previous analyses was applied with the participation of students from the experimental and control groups. Before starting the intervention procedure independent sample t-tests for each CD of the Octalysis were carried out and it was aimed at diagnosing situations in both groups. It was important to know about both groups so that it could be possible to compare and contrast the findings from pre-intervention and post-intervention period, and determine the effectiveness of the intervention.

In addition to the quantitative data, qualitative data were collected via semi-structured interviews both from the teachers and students in the pre-intervention period. It was aimed to illustrate Octalysis frameworks both for the general perception of language classes and writing classes. Students and teachers from both experimental and control groups were asked to evaluate their language classes in general and then for writing. Moreover, they were asked to grade their language learning environment in terms of eight CDs of Octalysis. Then, all the grades

obtained from interviews and means scores obtained from pre-intervention surveys were put together and new average scores were computed. These averages were used to draw illustrations of the Octalysis frameworks.

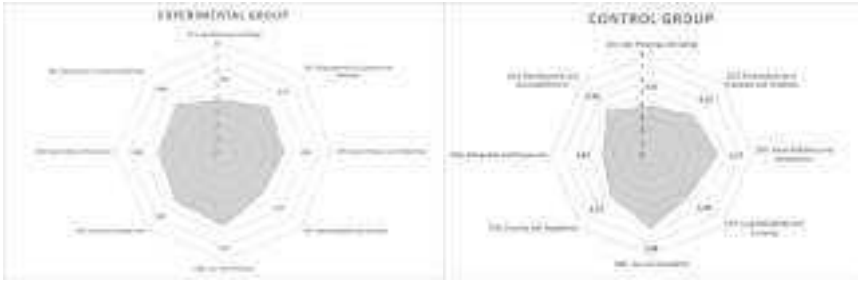


Figure 35. Comparison of pre-intervention general evaluation Octalysis of experimental and control groups

As can be understood from Figure 35 there was a balanced distribution of all CDs in both groups. The highest drive behind students' motivation in both groups were the avoidance, which meant that they were studying not to attain bigger and meaningful goals but to avoid failing in language classes, which was in line with Elliot (1999, 2006), Kaczmarek, Behnke, Enko, Kosakowski, Guzik and Hughes (2021), who state that when individual start to act to stay away the negative consequences then it is possible to mention about the existence of avoidance motivation. Although this kind of motivation is effective in short-term as suggested by Dörnyei (2021), it is highly unlikely to maintain the motivation in the face of challenges because if challenges result in avoidance then it could be said that they become a source of anxiety which, in turn, yield avoidance (Sunawan, Sugiyo & Pranoto, 2021; Henry & Davyenko, 2020). In this sense, it was possible to mention that there were serious problems regarding the type of goal attainment by students before the intervention. Expectedly, this could explain why the weakest drive was CD 1 in both groups which is found to be a requirement for the stimulation of intrinsic motivation when considered as a type of approach motivation (Daumiller & Zarrinabadi, 2021). Apart from the dichotomy of CD 1 and CD 8, other CDs were almost equally distributed in both groups.

In addition to the general evaluation of language environments, students and teachers from both groups were also asked to identify the skills that they thought to have biggest problems in. Findings showed that writing was perceived as one of the

most problematic skill. Therefore, other Octalysis frameworks were drawn for writing classes.

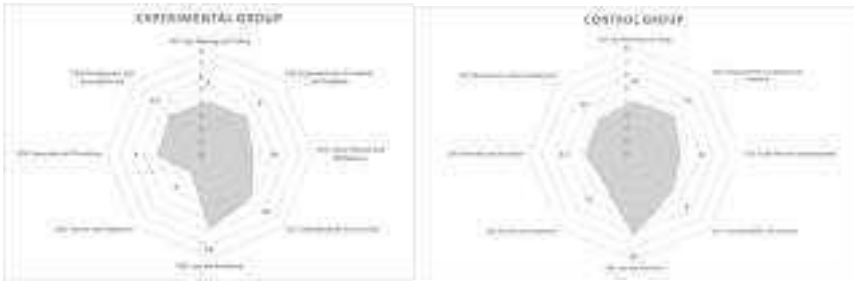


Figure 36. Comparison of pre-intervention writing evaluation Octalysis of experimental and control groups

Different from general evaluation, for writing classes even though there are similarities between the two groups, there was not a balanced distribution of CDs in either group. The power of CD 8: Loss and Avoidance was more than general evaluation classes and in turn, CD 1: Epic Winning and Meaning was the one the weakest CDs in both groups. That is, instead of pursuing big and meaningful goals in their writing classes, students were writing just to complete the tasks. According to the interview findings, the most important thing was not to fail in the class and not to lose marks due to incomplete tasks. However, in their interviews students also acknowledge that they wrote superficially and minimum was enough for most of them. This finding is in line with the literature. Louick and Muenks (2021) and Dörnyei (2021) point out that when students try to be autonomous and achieve bigger goals, they have an approach goal orientation whereas as they try to avoid poor performances, the avoidance motivation is in action.

Another biggest problem diagnosed by both students and teachers was the lack of interaction among students and engagement to the tasks which is a highly significant part of teaching and learning goals in language classrooms (Sert, 2015). Sert (2019) argue that provided that interactional activities among students are reinforced, then it becomes easier to provide better opportunities for learning to occur.

Interviews with students yielded enough evidence to state that the number of active students in each group was limited to only 5-6 students, which was a sign of lack of interaction among students, and the rest of the students were indifferent to the writing activities. In addition, even students wrote, they mostly did it individually

and they did not feel the need to interact with each other. However, Dörnyei and Ushioda (2011) state that there should be positive interaction between students to be able to mention about the existence of classroom level motivation. Based on their arguments it could be stated that low levels of interactions among students in classroom environment might be a factor affecting group dynamics and hinder students' motivation.

According to the interviews most of the students thought that it was impossible to have interactive writing classes in which they cooperate with their friends and teachers. In this sense this could be a problem. As Tiwari (2021) points out for the highest efficiency in language classrooms, a good amount of interaction needs to exist between students. In this sense, it could be said that these findings are consistent with the literature.

Research question 3: Can semi-structured pre-intervention interviews and pre-intervention survey provide data to draw Octalysis frameworks of experimental and control groups? Based on the above-mentioned findings, it could be said that it is possible to draw the illustrations of the Octalysis frameworks (Chou, 2015) for experimental and control groups by using the data which were obtained thorough semi-structured interviews and survey prior to the gamification intervention.

The illustrations of writing classes and interview transcriptions also revealed that the role of interesting applications and tasks in writing classes was another highlighted issue. Many students criticized their writing classes as being monotonous. According to them, writing classes lacked the elements to trigger their curiosity. Oxford (2016) and Silvia (2017) point out that curiosity contributes the development of positive language learning skills. Accordingly, it was argued that the monotonous nature of writing classes might be the cause of low engagement rates.

Another distinctive difference between both groups were about the role of CD 6: Scarcity and Impatience. In experimental group students felt the CD 6 had little effect on their writing motivation whereas students from control group stated relatively higher perception. Since in both groups there were similar circumstances in writing classes, this difference might stem from the applications of the teachers.

Research question 4: Is it possible to identify motivational drives by using pre-intervention Octalysis frameworks in experimental and control groups? The

illustrations of Octalysis frameworks in both groups showed that it was possible to identify the drives underlying students' language learning experiences (Chou, 2015). The illustrations provided a practical and simple way to diagnose drives (Duarte & Cruz, 2018) that led students act.

Research question 5: Are there any differences between the experimental and control groups before the implementation of gamification intervention? To sum up, findings of independent sample t-tests for each dimension of Octalysis Framework and transcriptions revealed no significant differences between control and experimental groups. In other words, these findings could be interpreted in the way that students in both classrooms perceived their language learning environment from similar point of views before the implementation of seven-week intervention.

Intervention. After diagnostic process, according to the findings Octalysis dashboard was drawn in the next step and a gamification intervention system was planned. In the subsequent seven weeks, intervention was applied and during this time quantitative data were collected in terms of the experience points students gained. In this way it was possible to keep track of students' engagement levels and interaction frequencies as well as their developments.

At the end of the intervention, there were 18 students who were still attending classes in experimental group. And all the students participated in writing activities with varying degrees. Although the end of the academic year was very close and significant demotivation levels was observed in control group, students in experimental group were still engaging into the writing activities. In this respect, it may not be wrong to assume that this intervention had positive effects on creating long-term motivation for students who previously had problems in participating writing classes.

These findings were in line with the literature. Dicheva, et al. (2015) state that there is a strong relationship between gamification and learning. Similarly, Xu, Lio, Dhaliwal, Andrei, Balakrishnan, Nagani and Samadder (2021) find out that by increasing social interaction, and utilizing the game elements of points, badges and leaderboards, gamification interventions improve intrinsic motivations which in turn may have a positive effect on learning. Dehghanzadeh, Fardanesh, Hatami, Talaee and Noroozi (2019) emphasize the possible positive outcomes of gamification in

terms of language learning activities. They state that since gamification is motivating, enjoyable, fun and engaging, it is likely to use gamification interventions to improve students' performances in language learning classes. The researchers who point out the positive outcomes of gamification interventions is not limited to these authors. Researchers such as Ruiz-Banuls, Gomez-Trigueros, Rovira-Collado and Rico-Gomez (2021), Laksanasut, Seubsang (2021), Bai, Hew and Huang (2020), Seixas et al., (2016), Hanus and Fox (2015), Caponetto et al., (2014), Lee and Hammer (2011), Chou (2015), Barata et al., (2017), Yıldırım and Demir (2014), Güler and Güler (2015) and Turan et al., (2016) are also mention about similar positive outcomes of such interventions.

As it was stated in strategy dashboard in pre-intervention stage, the primary goals of this intervention were to increase the social dimension (CD 5: Social Influence and Relatedness) in writing classes and enable students to enjoy their own learning experiences (CD 3: Empowerment of Creativity and Feedback) throughout the process by establishing strong feedback mechanisms. In addition, it was also aimed at decreasing the level of avoidance motivation (CD 8: Loss and Avoidance) which was creating short-term motivation and resulted in students' getting uninterested in writing classes.

Among secondary aims were increasing students writing competences (CD 2: Development and accomplishment) and feeling of ownership (CD 4: Ownership and Possession) in writing classes. While trying to ensure above-mentioned objectives, it was paid attention to integrate novel and interesting mediators (CD 6: Unpredictability and Curiosity) and make careful use of scarce and limited items (CD 6: Scarcity and Impatience) such as points, badges and leaderboards. Finally, it was expected to increase students' long-term motivation although there might be ups and downs at motivation levels throughout the process, which was theorized as Directed Motivational Currents by Dörnyei et al. (2016).

In addition, the data collected throughout the gamification intervention provided enough evidence to assume that designing gamification intervention based on Octalysis theory would promote long-term motivation in writing classes, which in line with the findings of Marisa, Ahmad, Maukar, Marcus and Widodo (2020) who support the idea that Octalysis framework can be used to develop long term motivation.

At the beginning the pre-intervention interview and survey findings showed there was no difference between the two groups. Teachers dealing with writing activities in both classes complained about the low engagement levels and low-quality writings. They stated that under normal circumstances the number of students who took part in writing activities was limited to only five or six students, which was lower than 30% percent of many classes. In addition, every year students experienced backwash effect of the final proficiency exam which would be used to determine their success in language preparation program, and which caused many students gave up in writing classes. In many classrooms towards the end of the year, there were no students who kept writing. Moreover, students in both groups disliked the individual and monotonous nature of their classes.

Findings collected throughout the intervention showed there was a significant change in experimental group. When students' performances in 7-week intervention were examined, on the average over 67% of the students took part in writing activities in experimental group whereas this rate was only 30% in control group, which meant the intervention positively affected the engagement rates in experimental group, which is consistent with the findings of Seifert and Gez (2021), Jonathan and Recard (2021), Bernik (2021), Madrid and de Jesus (2021), Zhang (2021), Palimbong (2019), Seixas et al., (2016), Hanus and Fox (2015), Caponetto et al., (2014). In addition, in experimental group although some students preferred to wait until they felt ready to take part in the activities, at the end of the term there were no students who had not participated in writing activities. But in control group, nine students had never participated in writing activities in any way, and in the final two weeks, none of the students in this group carry out any kind of writing activity because they started to study for the final exam. This finding proves the fact that it is not possible to sustain long-term motivation when avoidance motivation is so powerful as in the case of control group. This finding complies with the findings of Henry and DavyDenko (2020) who state that it is not possible to sustain long-term motivation when avoidance motivation is constant. Similarly, Kim and Castelli (2021) point out that gamification interventions have the potential to improve the learning outcomes. Both pre- and post-intervention findings showed that CD 8: Loss and Avoidance, which showed the power of avoidance motivation, was very powerful in control group, and their motivation diminished after some time, which is consistent

with the findings of Amini (2021) who states that avoidance motivation increases anxiety and this may have negative consequences regarding students' motivation levels. Similarly, Kantaridou, Machili and Papadopoulou (2021) argue that avoidance motivation result in lower language proficiency levels and demotivation. However, in experimental group post-intervention findings showed that DMC based Octalysis gamification intervention was successful in sustaining long-term motivation, which was reflected in the interaction and engagement ratings.

Post-Intervention. Research question 6: How is the DMC integrated Octalysis gamification intervention perceived by the participants in the experimental group? Findings of the 7-week gamification intervention and data from the post-intervention interviews showed that students and the teacher in experimental group perceived the gamification intervention positively and data analyses at these stages provided encouraging findings in terms of the effectiveness of gamification intervention in writing classes. These findings add to the existing literature. Lam, Hew and Chiu (2017), El Tantawi, Sadaf and Alhumaid (2018), Wiethof, Tavanapour and Bittner (2021) argue that gamification intervention significantly improves students' academic writing skills.

According to the teacher of experimental group there was a remarkable change in terms of students' writing aims. In addition, it was found out that systematic use of points, badges and leaderboards affected positively and the teacher thought that they helped students develop their competencies while use of points and badges did not resulted in a similar way in control group. Dicheva et al., (2015), Barata et al., (2017), Lister and College (2015) and Gåsland (2011) state that points, badges and leaderboard are among the most frequently used game elements. The finding of the current study supports the findings of the other researchers such as Hew et al. (2016), Goehle (2013), Charles et al. (2011), De-Marcos et al. (2014), Goehle (2013), Chou (2016), Chou (2019) who obtain a positive view about the use of points, badges and leaderboards in gamificaiton interventions.

Regarding the role of gamification intervention on CD 3, the teacher of experimental group clearly identified that students were enjoying in the process of writing in that 7-week period. The use of different feedback mechanics that aimed at enabling students interact with each other provide opportunities for weaker

students to get involved into the writing processes when they felt ready to do so. This was another distinctive indicator of positive effect of gamification intervention for the teacher of experimental group. Accordingly, the frequency of engagement in experimental group increased unprecedently with the inclusion of almost all students into the writing processes, which is in line with the findings of Looyestyn, Kernot, Boshoff, Ryan, Edney and Maher (2017) who conclude that gamification is an effective way of increasing engagement rates in online programs. On the contrary, the engagement frequencies dropped to zero probably due to negative backwash effect of final proficiency exam. Finally, the teacher of the experimental group indicated that many students started to hand in complete and original papers, which could be an indicator of a decrease in CD 8: Loss and Avoidance.

Like teacher's views, students in experimental group stated positive views in terms of the effectiveness of the gamification interview. Students form this group stated that throughout the process their aim for writing shifted and towards the end of the intervention they were believing they could strive for bigger goals. In addition, findings from the post-intervention interviews proved that gamification intervention helped students feel their writing competences improved significantly. Moreover, students from the experimental group highlighted the importance of how much they started to enjoy from writing activities which had been criticized as being highly monotonous and boring before. Accordingly, they thought that writing processes themselves became the source of fun and feedback mechanisms seemed to facilitate this.

Another primary aim that had been given importance in the design process of gamification intervention was to increase the interaction among students and their engagement frequencies. As understood from students' post-intervention comments, the elements integrated into the system had boosted the engagement and interaction levels in experimental group. Especially, the use of points system and design of the tasks were found to be highly effective in enabling less strong students took part in writing process. Finally, students in experimental group stated that their motivation for writing changed significantly compared to their pre-intervention motivations. As a result of the gamification intervention their fear of failure decreased and they did not need to write just to get mark, which was a sign of change in terms of the effect avoidance motivation.

Based on the above-mentioned findings, it could be stated that students in experimental group adopted positive points of views in terms of the effectiveness of the DMC integrated Octalysis gamification intervention, which is again in line with the findings of Güler and Güler (2015), Yıldırım and Demir (2014), Sheldon (2012), and Barata et al., (2013) who emphasize the positive influence of gamification interventions.

Having completed the 7-week gamification intervention in experimental group, students and teachers from both experimental and control groups were invited to evaluate their writing classes by grading them. By combining the data obtained from interviews, the following Octalysis frameworks for overall writing evaluation in control and experimental groups were drawn.

Figure 37. Post-intervention Octalysis frameworks

As can be seen from the above illustrations, there was almost no change in terms of students' the perceptions in control group for writing classes. As in the pre-intervention period, the most powerful drive behind their writing motivation was CD 8: Loss and Avoidance. On the contrary, an increase could be observed. This might be due to the approaching end-of-year proficiency exams. The data obtained from the post-intervention interviews supported this finding as students stated that final exam was much more important and there was no need to waste their efforts for writing. Also, the teacher who was responsible for the writing classes in control group indicated that in the last two weeks, no papers were handed in and the number of engagements in writing classes dropped to zero.

However, when the illustration of experimental group was investigated, it could be seen that there was a dramatic change in terms of the power of avoidance motivation. Before the intervention, CD 8 was very powerful in experimental group, too. But, after the implementation, it was seen that it turned into one of the weakest CDs in experimental group. Since quantitative and qualitative findings proved that

the circumstances had been similar in both groups before the intervention, it would not be wrong to assume that the change resulted from the intervention itself. The qualitative data from the interviews and quantitative data from the 7-week intervention period supported this finding. Accordingly, it could be stated that DMC embedded Octalysis gamification intervention was successful in motivating students to attain bigger and meaningful goals by reducing their avoidance motivation.

Research question 7: Are there any differences between the experimental and control groups after the intervention? Findings showed that there was a considerable difference between the experimental and control groups at the end of 7-week intervention. When pre- and post-intervention Octalysis frameworks were examined, it could be seen that there was little change in control group whereas the change in experimental group was apparent.

Primary research question: Can DMC integrated Octalysis as a gamification model be used in EFL classrooms to enable sustainable long-term motivation? According to Dörnyei et al. (2016) Directed Motivational Currents can be utilized at individual levels to canalize individuals to strive for their personal and clearly defined goals. However, Dörnyei et al. (2016) question the usability of DMCs at group level due to the atypical nature of the phenomenon. The findings of the current study provide evidence to their question. Quantitative and qualitative data analyses carried out within the context of this study showed that it was possible to use DMC embedded Octalysis gamification interventions in EFL classrooms to motivate groups of students. That is, the findings indicate that DMCs can operate at group levels and language learners can "find themselves in a collective state of motivational 'hyperdrive'" (Dörnyei et al., 2016, p. 141). In this respect, it was found out that it was possible to use gamification theories to design focused interventions to motivate the entire classes of learners together.

Pedagogical implications

Above-mentioned findings yielded some pedagogical implications which would help language learners, teachers, trainers and policy makers. By using Octalysis framework, it would be possible to diagnose what drives affect learners' motivation and plan suitable systems in accordance with the necessities.

One of the fuzzy points regarding the applicability of DMC is how to realize it in groups of language learners. For successful DMC group project, (Dörnyei et al., 2016) relevant and real goal content, tangible outcomes and clear success criteria need to be defined. Indeed, these requirements should be met for successful gamification designs. In this respect it could be stated that by using Octalysis dashboard (Chou, 2015) it was possible to clearly define the goals as in the form of desired actions; tangible outcomes as in the form of business metrics; and clear success criteria as in the form of win-states. In addition, within the context of group DMCs, various steps, sub-goals and progress checks need to be defined, which are also the basics of gamification interventions.

In group level DMCs, Dörnyei et al. (2016) state that goals need to be meaningful and interesting enough to capture learners' attention. In addition, learners need to feel a sense of ownership and control, and skill-challenge balance must be paid attention. Moreover, group dynamic should be appropriate because for successful DMCs to occur, there should be increased cooperation and agreement among group members, and they should be able to work towards common goal which should be sufficiently creative and engaging, and which are consisted of several sub-goals. There should also be frequent progress checks and regular feedback mechanisms to maintain the group DMCs. When Octalysis framework is investigated, it could be seen that it has similar principles with group level DMCs, and it is possible to utilize DMCs and Octalysis together. It is evident that eight CDs of Octalysis focus on the same aspects and requirements with DMCs.

The implications of this study suggest that by using Octalysis framework, teachers, trainers and policy makers can design systematic gamification interventions in language learning classrooms for groups of students. Designs made by using Octalysis Strategy dashboard can help teachers define the desired goals, tangible outcomes, feedback mechanics and clear success criteria. It can also be possible to decrease negative impacts of avoidance motivation by focusing on meaningful goals, and it might be possible to decrease the effects of avoidance approach motivation and create a sustainable environment in terms of long-term motivation. Besides, Octalysis framework can assist teachers manipulate their designs so that they can make use of creative activities more, arouse students interests and trigger their motivation. In addition, they can optimize social dimension

to create collaborative atmosphere and increase interaction and engagement rates in their classrooms.

Quantitative findings of the study indicate that it is possible to use Octalysis survey to make psychometric analyses in language education contexts which might help researchers interested in this field.

Limitations

Although this study provides many prospective benefits in terms of the use of gamification in language learning context, these benefits should be taken cautiously. First, data collected throughout the study were limited to students who had been attending in language preparation programs in several cities in Turkey.

Secondly, to measure the eight dimensions of the Octalysis framework, several scales and surveys had been selected and their related parts were adapted for the purposes of the current study. Therefore, in other studies there might be a need the evaluate the usability of the final scale in other contexts. Another, limitation related to the scale can be the number of items measuring the latent structures of Octalysis framework. Due to the relatively high number of latent variables, the number of items might be high and might cause some problems. As it is known that many multivariate statistical tests can be affected by the sample size. As the number of structures and the items increase, the sample size needs also be high enough to meet the preliminary assumptions. In this respect, especially in the first two EFA procedures, the sample size might not be so high. Future studies should take the size of the sample before starting the statistical procedures.

Another issue that might limit the generalizability of the findings was that all the participants who had taken part into the studies in scope of this study had the same proficiency levels. Thus, the generalizability of the findings was limited to this population. Further studies are needed to test the same procedures to evaluate the use of Octalysis.

In addition to participants of the second phase of the study were volunteered students and teachers from both control and experimental groups, and they had no previous knowledge about Octalysis framework and its eight dimensions. Although, they had been given a small briefing prior to the semi structured interviews, they

might have had difficulty in understanding the theory which might have also affected their evaluations. Further studies should be designed to overcome this problem and maybe more comprehensive training sessions or workshops can be planned before collecting interview data. Another limitation about the interviews was that owing to participants' lack of proficiency in the target language, all interviews were carried out in Turkish and then translated into English by the researcher, which might have an impact on the process.

Throughout this study gamification intervention was applied at group level to evaluate its role in group DMCs. Thus, although students' personal developments were recorded in experimental group, there was no record of such a thing in control group. That is, students' individual motivational levels did not taken into consideration. Maybe, in future studies, different designs can be planned to follow each student's motivational fluctuations in the long-term.

Finally, Octalysis (Chou, 2015) is a high sophisticated framework which consists of four different levels. However, in scope of the current study only level 1 Octalysis was applied. When taken individually, all levels of Octalysis framework incorporates different motivational and psychological theories ranging from Self-Determination Theory or Flow Theory to Dynamic Systems Theory. Therefore, in the future studies researchers should focus on different theories at different levels of Octalysis. Moreover, there are many variables that take role in gamification interventions such as game elements, game mechanics and so on. In this study, specific game elements such as points badges and leaderboards or game mechanics such as feedback mechanics were utilized. But their individual effects at did not paid attention. Further studies might be designed to evaluate their unique roles in gamification interventions.

References

Aarts, H., & Custers, R. (2012). Unconscious goal pursuit: nonconscious goal regulations and motivation. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 232-247). New York: Oxford University Press.

Abramovich, S., Schunn, C., & Higashi, R. M. (2013). Are badges useful in education?: it depends upon the type of badge and expertise of learner. *Educational Technology Research and Development*, 61, 217-232.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.

Aleven, V., Myers, E., Easterday, M., & Ogan, A. (2010). Toward a framework for the analysis and design of educational games. *The 3rd IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning*, (pp. 69-76).

Alsancak Sırakaya, D. (2017). Oyunlaştırılmış tersyüz sınıf modeline yönelik öğrenci görüşleri. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 36(1), 114-132.

Amini, D. (2021). Relationship between Persian as foreign language learners' anxiety, enjoyment and goal orientation. Journal of Teaching Persian to Speakers of Other Languages, 10(21), 75-100. doi: 10.30479/jtpsol.2021.15250.1525

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals, *Educational Psychologist*, 34:3, 169-189, DOI: 10.1207/s15326985ep3403_3

Antin, J. (2012). Gamification is not a dirty word. *Interactions*, 19(14).

Aparicio, A. F., Vela, F. L., Sanchez, J. L., & Montes, J. L. (2012). Analysis and application of gamification. *Proceedings of the 13th International Conference on Interacción Persona-Ordenador*, (p. 17).

Apostolou, M. (2013). Psychometric properties of the achievement goal questionnaire-revised administered to Greek university students. *Psychological Reports: Employment Psychology & Marketing*, 113(2), 380-403.

Armstrong, A. (2008). The fragility of group flow: the experiences of two small groups in a middle school mathematics classroom. *Journal of Mathematical Behavior*, 27(2), 101-115.

Arnold, F. (1910). *Attention and Interest: a study in psychology and education*. Macmillan.

Astaryan, V. S., Slevitch, L., Larzelere, R., Morosan, C., & Kwun, D. J. (2013). Effects of psychological ownership on students' commitment and satisfaction. *Journal of Hospitality & Tourism Education*, 25(4), 169-179.

Atkinson, J. W. (1964). *An Introduction to Motivation*. Princeton, NJ: Van Nostrand

Attali, Y., & Arieli-Attali, M. (2015). Gamification in assessment: do points affect test performance?, *Computers & Education*, 83, 57-63.

Baber, W. W. (2015). Course level gamification for higher education TEFL in Japan. «AGON» *Rivista Internazionale di Studi Culturali*, Linguistici e Letterari, 7.

Bai, S., Hew, K. F., & Huang, B. (2020) Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts, *Educational Research Review*, Vol. 30, ISSN 1747-938X

Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving student creativity with gamification and virtual worlds. *Proceedings of Gamification '13* (pp. 95-98). ACM Press.

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: a long-term study. *Computers in Human Behavior*, 71, 550-585.

Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, 47(4), 644-675.

Bartlett, M. S. (1954) A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society*, 16, 296-298.

Becker, K. (2021). What's the difference between gamification, serious games, educational games, and game-based learning? Academia Letters, Article 209. https://doi.org/10.20935/AL209

Bellotti, F., Berta, R., De Gloria, A., Lavagnino, E., Dagnino, M. F., Antonaci, A., & Ott, M. (2013). A gamified short course for promoting entrepreneurship among ICT engineering students, 2013 IEEE 13th International Conference on Advanced Learning Technologies, 31-32, doi: 10.1109/ICALT.2013.14.

Bentler, P. M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588–606.

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal*, 21(4), 755-765.

Berg, M. A. (2013). *Evaluating a game design for language learning: exploiting psychological needs to increase motivation for learning*. Norway.

Berkling, K., & Thomas, C. (2013). *Gamification of a software engineering course and a detailed analysis of the factors that lead to it's failure, 2013 International Conference on Interactive Collaborative Learning (ICL),* 525-530, doi: 10.1109/ICL.2013.6644642, 525-530.

Berlyne, D. (1960). *Conflict, arousal and curiosity*. New York: McGraw-Hill.

Bernik, A. (2021). Gamification framework for e-learning systems in higher education. *Tehnički Glasnik*, 15(2), 184-190. https://doi.org/10.31803/tg-20201008090615

Birant, K. U. (2014). Öğrenci oryantasyonunda web portalı temelli oyunlaştırma önerisi. *19."Türkiye'de İnternet" Konferansı Bildiriler Kitabı*, 91-98.

Bogost, I. (2011). Persuasive games: exploitation ware. Retrieved from Gamasutra: /http://www.gamasutra.com/view/feature/134735/persuasive_ga mes_exploitationware.php

Bollen, K. A., and Long, J. S. (1993). *Testing Structural Equation Models*. Sage, Newbury Park, CA.

Borges, S. S., Isotani, S., Durelli, V. S., Reis, H. M., Bittencourt, I. I., & Mizoguchi, R. (2017). Brazilian Portuguese cross-cultural adaptation and validation of the susceptibility to persuasion scale (Br-STPS)," 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), 2017, 72-73, doi: 10.1109/ICALT.2017.55..

Bozkurt, A., & Genç-Kumtepe, E. (2014). Oyunlaştırma, oyun felsefesi ve eğitim: gamification. *Akademik Bilişim*, 155-164.

Brathwaiter, B., & Schreiber, I. (2008). *Challenges for game designers*. Boston, Massachusetts: Charles River Media.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. (First Edition). NY: Guilford Publications, Inc.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K.A. & Long, J.S. [Eds.] *Testing structural equation models*. Newbury Park, CA: Sage, 136–162.

Buckingham, J. (2014). Open digital badges for the uninitiated. *The Electronic Journal for English as a Second Language*, 18(1).

Bunchball, I. (2010). *Gamification 101: an introduction to the use of game dynamics to influence behavior* (White paper). Retrieved from http://www.bunchball.com/sites/default/files/downloads/gamification101.pdf.

Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.

Caponetto, I., Earp, J., & Ott, M. (2014). Gamification and education: a Literature Review. In C. Busch (Ed.), Proceedings of the 8th European conference on games-based learning. London: Academic conferences and publishing international Limited.

Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables. In Bohrnstedt, G.W. & Borgatta, E.F. [Eds.] *Social measurement: current issues.* Beverly Hills: Sage.

Chan, L. (2014). Effects of an imagery training strategy on Chinese university students' possible second language selves and learning experiences. In K. Csizer, & M. Magid (Eds.), *The impact of self-concept on language learning.* Bristol, UK: Multilingual Matters.

Chang, E. Y. (2012). *Technoqueer: re/con/figuring*, Posthuman Narratives.

Chang, C., Liang, C., Chou, P., & Lin, G. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspectives from multimedia and media richness. *Computers in Human Behavior*. 71. 10.1016/j.chb.2017.01.031.

Charles, D., Charles, T., McNeill, M., Bustard, D., & Black, M. (2011). Game-based feedback for educational multi-user virtual environments. *British Journal of Educational Technology*, 42(4), 638-654.

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., & Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, 39, 216–236. https://doi.org/10.1007/s11031-014-9450-1.

Chen, H., Wigand, R. T., & Nilan, M. (2000). Exploring web users' optimal flow experiences. *Information Technology & People*, 13(4), 263-281.

Child, D. (2006). *The essentials of factor analysis*. Continuum, London.

Chou, Y.-K. (2015). *Actionable gamification: beyond points, badges, and leaderboards*. Octalysis Media.

Chou, Y. (2019). *Actionable gamification: beyond points, badges, and leaderboards*. Packt Publishing Ltd, ISBN 183921077X, 9781839210778

Cialdini, R. (2001). *Influence: science and practice* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Cialdini, R. B. (2012). The focus theory of normative conduct. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology,* 295–312. Sage Publications Ltd. https://doi.org/10.4135/9781446249222.n41

Civelek, M. E. (2017). *Yapısal eşitlik modellemesi metodoloj*isi. Beta yayıncılık, İstanbul.

Collier, Joel. (2020). *Applied structural equation modeling using AMOS: basic to advanced techniques*. 10.4324/9781003018414.

Corbett, S. (2010). Learning by playing: Video games in the classroom. Retrieved from: http://www.nytimes.com/2010/09/19/magazine/19video-t.html

Creswell, J. W., & Clark, V. L. (2011). *Designing and conducting mixed methods research*. SAGE.

Crookes, G., & Schmidt, R. W. (1991). Motivation: reopening the research agenda. *Language Learning*, 41, 469-512.

Cruz, C., Hanus, M., & Fox, J. (2015). The need to achieve: players' perceptions and uses of extrinsic meta-game reward systems for video game consoles. *Computers in Human Behavior*. 71. 10.1016/j.chb.2015.08.017.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass Publishers.

Csikszentmihalyi, M. (1990). *Flow: the psychology of optimal experience*. New York, NY: Harper and Row.

Csikszentmihalyi, M. (1996). *Creativity: flow and the psychology of discovery and invention*. New York: Haper Collins.

Csikszentmihalyi, M. (2008). *Flow: the psychology of optimal experience (Harper Perennial modern classics)*. Harper Perennial Modern Classics.

Cudeck, R., & O'Dell, L. L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin*. 115(3): 475–487, doi:10.1037/0033-2909.115.3.475.

Çağlar, Ş., & Kocadere, S. A. (2015). Çevirimiçi öğrenme ortamlarında oyunlaştırma. *Eğitim Bilimleri ve Uygulama*, 14(27), 83-102.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL Uygulamaları*. (II. Baskı). Ankara: Pegem Akademi.

Daumiller, M., & Zarrinabadi, N. (2021). My goal is to do the best that I can in this class: relevance of potential-based achievement goals for intrinsic motivation and course performance. *International Journal of Psychology*. https://doi.org/10.1002/ijop.12792

De Volder, M. L., & Lens, W. (1982). Academic achievement and future time perspective as a cognitive-motivational concept. *Journal of Personality and Social Psychology*, 42(3), 566-571.

Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuit: human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227-268.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125, 627-668.

Deci, E., & Ryan, D. (2012). *Overview of self-determination theory*. In The Oxford Handbook of Human Motivation, 85.

Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaee, E., & Noroozi, O. (2019). Using gamification to support learning English as a second language: a systematic review, Computer Assisted Language Learning, DOI: 10.1080/09588221.2019.1648298

De-Marcos, L., Dominguez, A., Saenz-de-Navarrete, J., & Pages, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education*, 75, 82-91.

Denny, P. (2013). The effect of virtual achievements on student engagement. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 763-772, ACM.

Deterding, S. (2011). Situated motivational affordances of game elements: A conceptual model. Gamification: using game design elements in non-gaming contexts, a workshop at CHI. Vancouver, BC: ACM.

Deterding, S. (2013). Gameful design for learning. *T+D*, 67(7), 60-63.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, 2425-2428.

De Vellis, R. (2003). *Scale development: theory and applications: theory and application*. Thousand Okas, CA: Sage.

Dewey, J. (1913). *Interest and effort in education*. Boston: Houghton Mifflin Co.

Di Bitonto, P., Corriero, N., Pesare, E., Rossano, V., & Roselli, T. (2014). Training and Learning in e-Health Using the Gamification Approach: The Trainer

Interaction. In C. Stephanidis, & M. Antona (Eds.), *Universal access in human-computer interaction: aging and assistive environment*, Vol. 8515, Springer.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: a systematic mapping study. *Educational Technology & Society*, 18(3), 75-88.

Dominguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: practical implications and outcomes. *Computers & Education*, 63(1), 380-392.

Donath, J. (2007). Signals in social supernets. *Journal of Computer-Mediated Communication*, 13(1), 231-251.

Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *Modern Language Journal*, 78, 273-284.

Dörnyei, Z. (2005). The *psychology of the language learner: individual differences in second language acquisition*. L. Erlbaum.

Dörnyei, Z. (2009). The L2 motivational self-system. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, language identity and the L2 self*, 9-42, Bristol: Multilingual Matters.

Dörnyei, Z. (2010). *Questionnaires in second language research: construction, administration, and processing*. 2nd Edition. London: Routledge.

Dörnyei, Z., & Chan, L. (2013). Motivation and vision: An analysis of future L2 self-images, sensory styles, and imagery capacity across two target languages. *Language Learning*, 63(3), 437-462.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching: motivation* (Second ed.). Harlow: Longman Pearson.

Dörnyei, Z., Henry, A., & Muir, C. (2016). *Motivational currents in language learning: frameworks for focused interventions*. New York & Abingdon, UK: Routledge.

Dörnyei, Z., Muir, C., & Ibrahim, Z. (2014). Directed Motivational Currents: Energising language learning through creating intense motivational

pathways. In D. Lasagabaster, A. Doiz, & J. M. Sierra (Eds.), *Motivation and foreign language learning: From theory to practice*, 9-29, Amsterdam: John Benjamins.

Duarte, S., & Cruz, M. (2018). The Gamification Octalysis *framework within the primary English teaching process: the quest for a transformative classroom*

El Tantawi, M., Sadaf, S., & AlHumaid, J. (2018), Using gamification to develop academic writing skills in dental undergraduate students. *Eur J Dent Educ*, 22, 15-22. https://doi.org/10.1111/eje.12238

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34, 149-169.

Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation and Emotion*, 30(2), 111–116, https://doi.org/10.1007/s11031-006-9028-7

Elliot, A. J. (2008). Approach and avoidance motivation and achievement goals. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation*, 3-14, New York: Psychology Press.

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: critique, illustration, and application. *Journal of Educational Psychology*, 100(3), 613-628.

Ellis, H., Heppell, S., Kirriemuir, J., Krotoski, A., & McFarlane, A. (2006). *Unlimited learning: Computer and video games in the learning landscape*. London: Entertainment and Leisure Software Publishers Association.

Erdoğdu, F., & Karataş, F. Ö. (2016). Fen eğitiminde oyunlaştırmanın farklı değişkenler üzerindeki etkilerinin incelenmesi. *Hoca Ahmet Yesevi Yılı Anısına Uluslararası Türk Dünyası Eğitim Bilimleri ve Sosyal Bilimler Kongresi*, 73-80.

Erümit, S. (2016). *OYUNLOyunlaştırma yaklaşımlarının eğiitimde kullanımı: tasarım tabanlı bir araştırma*. Erzurum.

Erten, İ. H. (2015). Validating my-self-as-a-learner scale (MALS) in the Turkish context. *Novitas-ROYAL* (Research on Youth and Language), 9(1), 46-59.

Farrell, A., & Rudd, J. M. (2009). Factor analysis and discriminant validity: a brief review of some practical issues. In Australia-New Zealand Marketing Academy Conference. Australia-New Zealand Marketing Academy.

Farzan, R., DiMicco, J. M., Millen, D. R., Dugan, C., Geyer, W., & Brownholtz, E. A. (2008). Results from deploying a participation incentive mechanism within the enterprise. *Results from deploying a participation incentive mechanism within the enterprise*, 563-572.

Finch, H. (2006). Comparison of the performance of varimax and promax rotations: factor structure recovery for dichotomous items. *Journal of Educational Measurement*, 43: 39-52. https://doi.org/10.1111/j.1745-3984.2006.00003.x

Fitz-Water, Z., Jhonson, D., Wyeth, P., & Tjondronegoro, D. (2017). Driven to drive? Investigating the effect of gamification on learner driver behavior, perceived motivation and user experience. *Computers in Human Behavior*, 71, 586-595.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. https://doi.org/10.2307/3151312

Foster, J. A., Sheridan, P. K., Irish, R., & Frost, G. S. (2012). Gamification as a strategy for promoting deeper investigation in a reverse engineering activity. Proceedings of the 2012 American Society for Engineering Education Conference, 5456.

Freitas, S. (2006). Using games and simulations for supporting learning. *Learning, Media and Technology*. 31. 343-358. 10.1080/17439880601021967

Gaggioli, A., Milani, N., Mazzoni, E., & Riva, G. (2011). Networked flow: a framework for understanding the dynamics of creative collaboration in educational and training settings. *The Open Education Journal*, 4(1), 41-49

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C., & McIntyre, P. (1993). A student's contributions to second-language learning. Part II: affective variables. *Language Teaching*, 26(11), 1-11.

Garson, D. (2010). Statnotes: Topics in multivariate analysis: factor analysis.

Garvey, C. (1990). *Play*. Cambridge, MA: Harvard University Press.

Gaskin, J. (2018). Citing claims, Gaskination's StatWiki, http://statwiki.kolobkreations.com/.

Gaskin, J. (2019). *Data screening, detecting univariate outliers*. http://statwiki.kolobkreations.com/index.php?title=Data_screening

Gåsland, M. M. (2011). Game mechanic-based e-learning: a case study (Master's thesis).

Gee, J. P. (2003). *What video games have to teach us about learning and literacy. New York*, Palgrave/Macmillan.

Gee, J. P. (2008). Learning and Games. In K. Salen (Ed.), *The ecology of games: connecting youth, games, and learning,* 21-40, Cambridge, MA: The MIT Press.

Gee, J. P. (2014). *What video games have to teach us about learning and literacy.* Second Edition. Macmillan.

George, D., & Mallery, P. (2003). *Using SPSS for Windows step by step: a simple guide and reference (4th ed.*). London: Pearson Education.

George, D., & Mallery, P. (2013). *IBM SPSS Statistics 19 Step by Step a Simple Guide and Reference.*

Ghanizadeh, A., & Jahedizadeh, S. (2017). Directed motivational currents: the implementation of the dynamic web-based Persian scale among Iranian EFL learners. *Journal of Teaching Language Skills* (JTLS), 36(1), 27-56.

Giannetto, D., Chao, J., & Fontana, A. (2013). Gamification in a social learning environment. *Informing Science and Information Technology*, 10, 195-207.

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of grounded theory: strategies for qualitative research*. New York: Aldine De Gruyter.

Glen, S. (2013). Standardized residuals in statistics: what are they? From StatisticsHowTo.com: Elementary Statistics for the rest of us! https://www.statisticshowto.com/what-is-a-standardized-residuals/

Glover, I., Campell, A., Latif, F., Norris, L., Toner, J., & Tse, C. (2012). A Tale of one city: intra-institutional variations in migrating VLE platform. *Research in Learning Technology*, 20. DOI: 10.3402/rlt.v20i0.19190

Goehle, G. (2013). Gamification and web-based homework. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 23(3), 234-246.

Gökkaya, Z. (2014). Yetişkin eğitiminde yeni bir yaklaşım: Oyunlaştırma. *Hasan Âli Yücel Eğitim Fakültesi Dergisi*, 11(1), 71-84.

Grolnick, W. S., & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology*, 81(2), 143.

Groth, F. G. (2012). Gamification: State of the art definition and utilization. *Proceedings of the 4th seminar on Research Trends in Media Informatics*, 39-46.

Gulliksen, H., & Tukey, J. W. (1958). Reliability for the law of comparative judgment. *Psychometrika*, 23, 95–110.

Güler, C., & Güler, E. (2015). Çevrimiçi öğrenme ortamlarında oyunlaştırma: rozet kullanımı. *Eğitim ve Öğretim Araştırmaları Dergisi*, 4(3), 125-130.

Habgood, M. P. (2007). The effective integration of digital games and learning content. Unpublished doctoral dissertation. University of Nottingham, England.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2013). *Multivariate Data Analysis*. Pearson Education Limited

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? –a literature review of empirical studies on gamification. *The 47th Hawaii International Conference on System Sciences* (pp. 3025-3034). Hawaii: IEEE.

Hamari, J. (2013). Transforming homo economicus into homo-ludens: a field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications*, 12(4), 236-245.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80, 152-161.

Hardesty, D. M., & Bearden, W. O. (2004). The use of expert judges in scale development Implications for improving face validity of measures of unobservable constructs. *Journal of Business Research*, 57 (2004) 98 – 107. doi:10.1016/S0148-2963(01)00295-8

Harkness, J., & Schoua-Glusberg, A. (1998). Questionnaires in Translation. *ZUMA-Nachrichten Spezial,* 87-126.

Hefferon, K., & Boniwell, I. (2011). *Positive psychology: Theory, research and applications.* Maidenhead, UK: Open University Press.

Henry, A., & Davydenko, S. (2020), Thriving? or Surviving? An approach–avoidance perspective on adult language learners' motivation. *The Modern Language Journal*, 104: 363-380. https://doi.org/10.1111/modl.12635

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A New criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1): 115-135.

Hew, K. F., Huang, B., Chu, K. S., & Chiu, D. K. (2016). Engaging Asian students through game mechanics: Findings from two experiment studies. *Computers & Education*, 92-93, 221-236.

Hidi, S. (2001). Interest, reading and learning: Theoretical and practical considerations. *Educational Psychology Review*, 13, 191-210.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94(3), 319-340.

Holye, R. H. (1995). *Structural equation modeling: concepts, issues, and applications 1st edition*. SAGE Publications, Inc

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modeling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*. 7 (2), 191-205.

Houtari, K., & Hamari, J. (2012). Defining gamification: A service marketing perspective. Proceeding of the 16th International Academic MindTrek Conference, 17-22.

Hsu, C. L., & Lu, H. P. (2004). Why do people play on-line games? an extended tam with social influences and flow experience. *Information & Management*, 41(7), 853-868.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.

Hu, P., & Zhang, J. (2017). A pathway to learner autonomy: a self-determination theory perspective. *Journal of Asia Pacific Education Review*, 18(1), 147-157.

Huang, H. W., & Soman, D. (2013). A practitioner's guide to gamification of education. research report series: behavioral economics in action. University of Toronto –Rotman School of Management.

Hull, C. L. (1943). *Principles of behavior: an introduction to behavior theory*. New York: Appleton-Century-Crofts.

Huta, V., & Ryan, R. M. (2010). Pursuing pleasure or virtue: The differential and overlapping well-being benefits of hedonic and eudaimonic motives. *Journal of Happiness Studies*, 11(6), 735-762.

Huta, V., & Waterman, A. S. (2013). Eudaimonia and its distinction from hedonia: developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies*, 15(6), 1425-1456.

Jackson, S. A. (2000). Joy, fun and flow state in sport. In Y. L. Hanin (Ed.), *Emotions in sport*,135-155. Champaign, IL: Human Kinetics.

Jackson, S. A. (2012). Flow. In R. Ryan (Ed.), *The oxford handbook of human motivation* 127. OUP USA.

Jackson, S. A., & Csikszentmihalyi, M. (1999). *Flow in sports*. Champaign, IL: Human Kinetics.

Jenkins, P. (2015). *Gamification is the key to mobile customer engagement*.

Johnson, L., Becker, A., Estrada, V., & Freeman, A. (2014). NMC horizon report: 2014 higher education edition. Austin, Texas: The New Media Consortium.

Johnson, D., & Wiles, J. (2003). Effective affective user interface design in games. *Ergonomics*, 46(13-14), 1332–1345.

Jonathan, F., & Recard, M. (2021). The implementation of gamification concept inside online classroom activities to promote students' engagement. *International Journal of Indonesian Education and Teaching*, 5(2), 176-184. doi:https://doi.org/10.24071/ijiet.v5i2.3461

Jöreskog, K. G., & Sörbom, D. (1993). *Lisrel 8: Structural Equation Modeling with the Simplis Command Language*. Lincolnwood: Scientific Software International, Inc.

Julkunen, K. (1989). *Situation and task-specific motivation in foreign-language learning and teaching*. Joensuu. University Joensuu.

Jung, J. H., Schneider, C., & Valacich, J. (2010). Enhancing the motivational affordance of information systems: the effects of real-time performance feedback and goal setting in group collaboration environments. *Management Science*, 56(4), 724-742.

Juul, J. (2003). The game, the player, the world: looking for a heart of gameness. *Level Up: Digital Games Research Proceedings*, (pp. 30-45). UtrechtUniversity, Utrecht.

Kaczmarek, L. D., Behnke, M., Enko, J., Kosakowski, M., Guzik, P., & Hughes, B. M. (2021). Splitting the affective atom: divergence of valence and approach-avoidance motivation during a dynamic emotional experience. Current Psychology 40, 3272–3283, https://doi.org/10.1007/s12144-019-00264-3

Kaiser, H. F. (1974) An index of factorial simplicity. *Psychometrika*, 39, 31-36. http://dx.doi.org/10.1007/BF02291575

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. San Francisco: CA: John Wiley & Sons.

Kaptein, M., Ruyter, B., Markopoulos, P., & Aarts, E. (2012). Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems*, 1025.

Karageorghis, C. I., Vlachopoulos, S. P., & Terry, P. C. (2000). Latent variable modelling of the relationship between flow and exercise-induced feelings: an intuitive appraisal perspective. *European Physical Education Review*, 6, 230-248.

Karataş, E. (2014). Eğitimde oyunlaştırma: Araştırma eğilimleri. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 15(2).

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. H. (2009). The Curiosity and exploration inventory-ii: development, factor structure, and psychometrics. *Journal of Research in Personality*, 43(6), 987-998.

Kashdan, T. B., McKnight, P. E., Fincham, F. D., & Rose, P. (2011). When curiosity breeds intimacy: taking advantage of intimacy opportunities and transforming boring conversations. *Journal of Personality*, 79(6), 1369-1402.

Kauffman, D. F., & Husman, J. (2004). Effects of time perspective on student motivation: introduction to a special issue. *Educational Psychology Review*, 16(1), 1-7.

Kim, A. J. (2011). *Gamification 101: Designing the Player Journey*.

Kim, J., & Castelli D. M. (2021). Effects of gamification on behavioral change in education: a meta-analysis, *International Journal of Environmental Research and Public Health* 18, no. 7: 3550. https://doi.org/10.3390/ijerph18073550

Kim, T. Y. (2009). Korean elementary school students' perceptual learning style, ideal L2 self, and motivated behavior. *Korean Journal of English Language and Linguistics,* 9(3), 261-286.

Kirriemuir, J., & McFarlane, A. (2004). Literature review in games and learning (Research report No. 8). A NESTA Futurelab Research report.

Klaff, O. (2011). *Pitch anything: an innovative method for presenting, persuading and winning the deal*. New York: McGraw-Hill.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

Kline, R. B. (2011). *Principles and practice of structural equation modelling* (3rd ed.). New York, NY: Guilford publications.

Klopfer, E., Osterweil, S., & Salen, K. (2009). *Moving learning games forward: The education arcade*. Cambridge, MA: Massachusetts Institute of Technology.

Kocadere, S. A., & Çağlar, Ş. (2015). The design and implementation of a gamified assessment. *Journal of e-Learning and Knowledge Society*, 11(3), 85-99.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, 35, 179-188.

Koster, R. (2004). *A Theory of fun for game design*. Scottsdale, AZ: Paraglyph Press.

Krapp, A. (2002). An educational–psychological theory of interest and its relation to SDT. In E. L. Deci, & M. L. Ryan (Eds.), *Handbook of self-determination research*, 405-427. New York, NY: University of Rochester Press.

Kuo, M. S., & Chuang, T. Y. (2016). How gamification motivates visits and engagement for online academic dissemination: an empirical study. *Computers in Human Behavior*, 55, 16-27.

Laksanasut, S., & Seubsang, P. (2021). Effects of gamified learning intervention on Thai students' motivation for learning English. Book Chapters of the 1st Jakarta International Conference on Social Sciences and Humanities

Lam, L. W. (2012). Impact of competitiveness on salespeople's commitment and performance. *Journal of Business Research*, 65(9), 1328-1334.

Lam, Y. W., Hew, K. F., & Chiu, K. F. (2017). Improving argumentative writing: Effects of a blended learning approach and gamification. *Language Learning & Technology*, 22(1), 97–118. https://dx.doi.org/10125/44583

Landers, R. N., & Armstrong, M. B. (2017). Enhancing instructional outcomes with gamification: an empirical test of the technology-enhanced training effectiveness model. *Computers in Human Behavior*, 71, 499-507.

Landers, R. N., & Callan, R. C. (2011). Casual social games as serious games: the psychology of gamification in undergraduate education and employee training. In M. Ma, A. Oikonomou, & L. C. Jain (Eds.), *Serious games and edutainment applications*, 399-423. London: Springer.

Landers, R. N., Bauer, K. N., & Callan, R. C. (2017). Gamification of task performance with leaderboards: a goal setting experiment. *Computers in Human Behavior*, 71, 508-515.

Lawley, D. N., & Maxwell, A. E. (1963). *Factor Analysis as a Statistical Method*. London: Butterworth.

Lazzaro, N. (2004). Why we play games: four keys to more emotion without story. Game Dev Conf.

Lee, J. J., & Hammer, J. (2011). Gamification in education: what, how, why bother? *Academic Exchange Quarterly*, 15(2), 1-5.

Li, W., Grossman, T., & Fitzmaurice, G. (2012). GamiCAD: a gamified tutorial system for first time AutoCAD users. *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, 103-112. Cambridge, MA: ACM.

Lillemor, R. M., & Hallberg, M. (2009). *The "core category" of grounded theory: making constant comparisons*, 141-148

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.

Lister, M. C., & College, H. (2015). Gamification: The effect on student motivation and performance at the post-secondary level. *Issues and Trends in Educational Technology*, 3(2).

Locke, E., & Latham, G. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.

Looyestyn, J., Kernot, J., Boshoff, K., Ryan, J., Edney, S., & Maher, C. (2017). Does gamification increase engagement with online programs? A systematic review. *PLOS ONE* 12(3): e0173403. https://doi.org/10.1371/journal.pone.0173403

Louick, R., & Muenks, K. (2021). *Leveraging motivation theory for research and practice with students with learning disabilities, theory into Practice*, DOI: 10.1080/00405841.2021.1932154

Macay, J. (2014). Applications and implications of the l2motivational self system in a catalan efl context. In K. Csizer, & Magid M. (Eds.), *The impact of self-concept on language learning*, 377-400. Bristol, UK: Multilingual Matters.

Madrid, M. A. C., & de Jesus, D. M. A. (2021) Towards the design and development of an adaptive gamified task management web application to increase student engagement in online learning. In: Sottilare R.A., Schwarz J. (eds) *Adaptive Instructional Systems. Design and Evaluation*. HCII 2021. Lecture Notes in Computer Science, vol. 12792. Springer, Cham. https://doi.org/10.1007/978-3-030-77857-6_14

Magid, M. (2014). A motivational program for learners of English: an application of the l2 motivational self-system. In K. Csizer, & M. Magid (Eds.), *The impact of self-concept on language learning,* 333-356. Bristol, UK: Multilingual Matters.

Malhotra, N. K., Dash, S. (2011). *Marketing Research an Applied Orientation*. London: Pearson Publishing.

Marisa, F., Ahmad, S. S. S., & Maukar, A. L., Marcus, R. D., and Widodo, A. A (2020). Evaluation of student core drives on e-learning during the covid-19 with octalysis gamification framework. *International Journal of Advanced Computer Science and Applications*, Vol. 11, No. 11, 104-11

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41(9), 954-969.

Markus, H., & Ruvolo, A. (1989). Possible selves: personalized representations of goals. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology*, 211-241. Hillsdale, NJ: Erlbaum.

Maslow, A. H. (1968). *Toward a psychology of being*. New York: Van Nostrand.

McGonigal, J. (2011). *Reality is Broken* (Vols. New York, NY). Penguin Group.

McGonigal, J. (2010). Gaming can make a better world. Retrieved 12 07, 2017, from TED:https://www.ted.com/talks/jane_mcgonigal_gaming_can_make_a_better_world/details

McKnight, P. E., McKnight, K. M., & Figueredo, A. J. (2007). *Missing data: A gentle introduction*. Guilford Press

Meşe, C. (2016). *Harmanlanmış öğrenme ortamlarında oyunlaştırma bileşenlerinin etkinliği.* Eskişehir.

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013). Disassembling gamifica-tion: the effects of points and meaning on user motivation and performance. *CHI'13 Extended Abstracts on Human Factors in Computing Systems,* (pp. 1137–1142). Paris.

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, 71, 525-534.

Meydan, C. H., & Şeşen, H. (2015). *Yapısal eşitlik modellemesi amos uygulamaları*. Ankara: Seçkin Yayınevi.

Meyer, K. (2008). Do rewards shape online discussions? *Journal of Interactive Online Learning*. 7, 2, ISSN: 1541-4914.

Michael, D. R., & Chen, S. (2006). *Serious games: games that educate, train and inform*. Thomson Course Technology.

Muir, C. (2016). The dynamics of intense long-term motivation in language learning: Directed Motivational Currents in theory and practice. Doctoral Thesis.

Muntean, C. I. (2011). Raising engagement in e-learning through gamification. The 6th International Conference on Virtual.

Nacke, L. E., & Deterding, S. (2017). The maturing of gamification research. *Computers in Human Behavior*, 71, 450-454.

Nakamura, J., & Csikszentmihalyi, M. (2002). The concept of flow. In C. R. Synder, & S. J. Lopez (Eds.), *Handbook of positive psychology*, 89-105. New York: Oxford University Press.

Nakamura, J., & Csikszentmihalyi, M. (2009). The Concept of Flow. In C. R. Snyder, & S. J. Lopez (Eds.), *Oxford handbook of positive psychology*, 89-105. Oxford University Press, USA.

Nicholson, S. (2012). A User-Centered Theoretical Framework for Meaningful Gamification. *Proceedings of Games Learning Society 8.0 (GLS 8.0)*, 223-230). ETC Press.

Nunnally, J., & Bernstein, L. (1994). *Psychometric theory*. New York: McGraw-Hill Higher, INC.

Osborne, J. W., Costello, A. B., & Kellow, J. T. (2008). Best practices in exploratory factor analysis. In J. W. Osborne (Ed.), *Best Practices in Quantitative Methods*, 205-213. Thousand Oaks, CA: Sage Publishing.

Oxford, R. L. (2016). Toward a psychology of well-being for language learners: the 'EMPATHICS' vision. In P. D. McIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA*, 10-90. Bristol: Multilingual Matters.

Oxford, R., & Shearin, J. (1994). Language learning motivation: expanding the theoretical framework. *The Modern Language Journal*, 78(1), 12-28.

Pallant, J. (2011). *SPSS survival manual: a step by step guide to data analysis using the SPSS program*. 4th Edition, Allen & Unwin, Berkshire.

Palimbong, T. S. S. (2019). Octalysis gamification framework for enhancing students' engagement in language learning. Bachelor thesis, Universitas Pelita Harapan.

Papadopoulou, I., Machili, I., & Kantaridou, Z. (2021). 9 Task-specific strategy use in video-mediated integrated writing: the Greek EAP context. In Z. Gavriilidou & L. Mitits (Ed.), *Situating Language Learning Strategy Use: Present Issues and Future Trends*, 165-180. Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781788926720-014

Papies, E. K., & Aarts, H. (2010). Nonconscious self-regulation, or the automatic pilot of human behavior. In K. D. Vohs, & R. F. B. (Eds.), *Handbook of self-regulation: research, theory, and applications*, 125-142. New York: Guildford Press.

Peterson, C. (2006). *A primer in positive psychology*. New York: Oxford University Press.

Pilke, E. M. (2004). Flow Experiences in Information technology use. *International Journal of Human-Computer Studies*, 61(3), 347-357.

Prensky, M. (2001). *Digital game-based learning: fun, play and games. what makes games engaging*? NewYork: McGraw-Hill.

Rao, V. (2013). *Challenges of implementing gamification for behavior change: Lessons learned from the design of Blues Buddies*. Paris, France.

Rapp, A. (2017). Designing interactive systems through a game lens: An ethnographic approach. *Computers in Human Behavior*, 71, 455-468.

Raynor, J. O. (1974). Motivation and career striving. In J. W. Atkinson, & J. O. Raynor (Eds.), *Motivation and achievement* 369-387. Washington, DC: Winston & Sons.

Raynor, J. O., & Entin, E. E. (1983). The function of future orientation as a determinant of human behavior in step-path theory of action. *International Journal of Psychology*, 18, 463-487.

Reeve, J., Deci, E. L., & Ryan, R. M. (2004). Self-Determination theory: a dialectical framework for understanding socio-cultural influences on student motivation. In D. M. Mclnerney, & S. Van Etten (Eds.), *Big Theories Revisited*, pp. 31-60. Greenwich, CT: Information Age Press.

Reeves, B., & Read, J. L. (2009). *Total engagement: using games and virtual worlds to change the way people work and businesses compete*. Boston: Harvard Business School Press.

Riani, M., Torti, F., & Zani, S. (2011). Outliers and Robustness for Ordinal Data. In: Kennett R.S. and Salini S. (eds.), *Modern Analysis of Customer Surveys: With Applications Using R*, 155-169, Wiley, New York. ISBN 9780470971284, DOI 10.1002/9781119961154.

Ringdon, E. E. (1998). Structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research*, 251-294. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Robertson, M. (2010). Can't play, won't play. Retrieved from Hide & Seek: /http://hideandseek.net/2010/10/06/cant-play-wont-play/S (accessed10.10.14).

Ruiz-Bañuls, M., Gómez-Trigueros, I., Collado, J. R. & Rico-Gómez, M. L. (2021). Gamification and transmedia in interdisciplinary contexts: a didactic intervention for the primary school classroom. *Heliyon*. 7. 10.1016/j.heliyon.2021.e07374.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic, social development, and well-being. *American Psychologist*, 55(1), 68-78.

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: motivation, learning, and well-being. In K. R. Wentzel, & A. Wingfield (Eds.), *Handbook of school motivation*. New York: Routledge.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivation pull of video games: a self-determination theory approach. *Motivation and Emotions*, 30(4), 347-363.

Sailer, M., Hense, J., Mandl, H., & Klevers, M. (2013). Psychological perspectives on motivation through gamification. *Interaction Design and Architecture(s) Journal*, 19, 28-37.

Sakamoto, M., Nakajima, T., & Alexandrova, T. (2012). Value-Based design for gamifying daily activities. *Proceedings of the 11th international conference on entertainment computing* (ICEC'12) (pp. 421-424). Berlin, Heidelberg: Springer- Verlag.

Salanova, M., Rodríguez-Sánchez, A. M., Schaufeli, W. B., & Cifre, E. (2014). Flowing together: a longitudinal study of collective efficacy and collective flow among workgroups. *The Journal of Psychology*, 148(4), 435-455.

Salen, K., & Zimmerman, E. (2004). *Rules of Play*. Cambridge, MA: MITPress.

Sampson, R. (2012). The language-learning self, self-enhancement activities, and self-perceptual change. *Language Teaching Research*, 16, 317-335.

Sarı, A., & Altun, T. (2016). Oyunlaştırma yöntemi ile işlenen bilgisayar derslerinin etkililiğine yönelik öğrenci görüşlerinin incelenmesi. *Turkish Journal of Computer and Mathematics Education*, 7(3), 553-567.

Sawyer, R. K. (2003). *Group creativity: music, theatre, collaboration*. Mahwah, NJ: Lawrence Erlbaum.

Sawyer, R. K. (2006). Group creativity: Musical performance and collaboration. *Psychology of Music*, 34(2), 148-165.

Scheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal wellbeing: the self-concordance model. *Journal of Personality and Social Psychology*, 76(3), 482-497.

Schell, J. (2008). *The art of game design: a book of lenses*. Morgan Kaufmann.

Schell, J. (2010). Visions of the gamepocalypse. Presentation. San Francisco, CA: Long Now Foundation.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: a survey. *International Journal of Human-Computer Studies*, 74, 14-31.

Seifert, T., & Gez, Y. (2021). Engaging students through escape games and a play environment during recess. *International Journal of Game-Based Learning*, 11(3). DOI: 10.4018/IJGBL.2021070101

Seixas, L. D., Gomes, A. S., & de Melo Filho, I. J. (2016). Effectiveness of gamification in the engagement of students. *Computers in Human Behavior*, 58, 48-63.

Selçuk, Ö., & Erten, İ. H. (2017). A display of patterns of changes in learners' motivation: dynamics perspective. *Novitas-ROYAL (Research on Youth and Language)*, 11(2), 128-141.

Seligman, M. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*.

Seligman, M. P., & Csikzentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, 55, 5-14.

Sert, O. (2015). *Social interaction and L2 classroom discourse*. Edinburgh University Press, Edinburgh (2015)

Sert, O. (2019). Classroom interaction and language teacher education. Walsh, S. & Mann, S. (Eds) *The Routledge handbook of English language teacher education*, 216-238. New York: Routledge

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal wellbeing: the self-concordance model. *Journal of Personality and Social Psychology*, 76(3), 482-497.

Sheldon, L. (2012). *The multiplayer classroom: designing coursework as a game*. Boston, MA: Cengage Learning.

Shernoff, D. J. (2013). *Optimal learning environments to promote student engagement.* Springer.

Shernoff, D. J., & Csikszentmihalyi, M. (2009). Flow in schools. In E. S. Heubner, & M. Furlong (Eds.), Handbook of positive psychology in *schools*, 131-145. Routledge.

Shernoff, D. J., Abdi, B., Anderson, B., & Csikszentmihalyi, M. (2014). Flow in schools revisited: Cultivating engaged learners and optimal learning environments. In M. Furlong, R. Gilman, & S. Heubner (Eds.), *Handbook of positive psychology in the schools (2nd edition)*, 211-226. Routledge, Taylor & Francis.

Silvia, P. J. (2017). Curiosity. In P. A. O'Keefe & J. M. Harackiewicz (Eds.), *The Science of Interest*, 97–107. Cham, Switzerland: Springer.

Simoes, J., Redondo, R. D., & Vilas, A. F. (2013). A social gamification framework for a K-6 learning platform. *Computers in Human Behavior*, 29(2), 345-353.

Skehan, P. (1989). *Individual differences in second language learning*. London: Arnold.

Song, M., & Zhang, S. (2008). EFM: A model for educational game design. In International Conference on Technologies for e-Learning and Digital Entertainment (pp. 509–517). Springer, Berlin, Heidelberg.

Sposito, V. A., Hand, M. L., & Skarpness, B. (1983). On the efficiency of using the sample kurtosis in selecting optimal lpestimators. *Communications in Statistics-simulation and Computation*, 12(3), 265-272.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, 42, 893-898.

Steinmetz, H. (2016). Re: Is it possible to continue with a low average variance extracted (AVE) value?. Retrieved from: https://www.researchgate.net/post/Is_it_possible_to_continue_with_a_low_average_variance_extracted_AVE_value/569361715cd9e38ec48b4596/citation/download

Stevens, J. (1992). *Applied multivariate statistics for the social sciences (2nd ed.)*. Lawrence Erlbaum Associates, Inc

Stott, A., & Neustaedter, C. (2013). *Analysis of gamification in education*.

Sunawan, S., & Pranoto, Y. K. S. (2021). Achievement goals and extraneous load predict germane load: The mediating effects of achievement emotions. *Malaysian Journal of Learning and Instruction*, 18(2), 215-234. https://doi.org/10.32890/mjli2021.18.2.8

Sümer, N. (2000). Yapısal eşitlik modelleri. *Türk Psikoloji Yazıları*, 3 (6), 49-74.

Şahin, M., & Samur, Y. (2017). Dijital çağda bir öğretim yöntemi: oyunlaştırma. *Journal of Ege Education Technologies*, 1(1), 1-27.

Tabachnick, B. G., & and Fidell, L. S. (2013). *Using Multivariate Statistics (sixth ed.)* Pearson, Boston (2013)

Teo, T., Tsai, L. T., & Yang, C.-C. (2013). Applying structural equation modeling (sem) in educational research: an introduction. In M. S. Khine (Ed.), Application of structural equation modeling in educational research and practice, 3-22. Boston: Sense Publishers.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC: American Psychological Association. ISBN: 1-59147-093-5).

Tiwari, T. D. (2021). Classroom interaction in communicative language teaching of public secondary schools in Nepal. *Indonesian Journal of English Language Teaching and Applied Linguistics,* 5(2), 373-386

Tunga, Y., & İnceoğlu, M. M. (2016). Oyunlaştırma tasarımı. 3rd. International Conference on New Trends in Education, 267-279).

Turan, Z., Avinc, Z., Kara, K., & Göktaş, Y. (2016). Gamification and education: achievements, cognitive loads, and views of students. *International Journal of Emerging Technologies in Learning,* 11(7), 64-69.

Ulicsak, M., & Wright, M. (2010). Games in education: serious games. *A Futurelab literature review*. Retrieved from Futurelab: www.futurelab.org.uk

Ullman, J. B. (2013). Structural equation modeling (chapter 14). In B. Tabachnick, & Fidell (Eds.), *Using Multivariate Statistics*, 681-785. Boston, MA: Pearson Education INC.

Ushioda, E. (1994). L2 motivation as a qualitative construct. *Teanga*, 14, 76-84.

Ushioda, E. (2008). Motivation and good language learners. In C. Griffits (Ed.), Lessons from good language learners, 19-34. Cambridge, UK: Cambridge University Press.

Van Dyne, L., & Pierce, J. L. (2004). Psychological ownership and feelings of possession: three field studies predicting employee attitudes and organizational citizenship behavior. *Journal of Organizational Behavior*, 25, 439-459.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review*, 41(2).

Van Meerkerk, E. (2017). Teacher logbooks and professional development: a tool for assessing transformative learning processes. *International Journal of Qualitative Methods*, 16, 1-8.

Voerman, L., Meijer, P. C., Korthagen, F. A., & Simons, R. J. (2012). Types and frequencies of feedback interventions in classroom interaction in secondary education. *Teaching and Teacher Education*, 28(8), 1107-1115.

Waterman, A. S. (1993). Two conceptions of happiness: contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, 64(4), 678-691.

Waterman, A. S. (2008). Reconsidering happiness: a eudaimonist's perspective. *The Journal of Positive Psychology*, 3(4), 234-252.

Werbach, K. (2014). (Re) Defining Gamification: A Process Approach. In A. Spagnolli, & L. Gamberini (Eds.), *Persuasive technology*, 266-272. Padova: Springer International Publishing.

Werbach, K., & Hunter, D. (2012). *For the win. how game thinking can revolutionize your business.* Philadelphia, PA: Wharton Digital Press.

Wiethof, C., Tavanapour, N., & Bittner, E. (2021). Designing and evaluating a collaborative writing process with gamification elements: toward a framework for gamifying collaboration processes. *AIS Transactions on Human-Computer Interaction*, 13(1), 38-61. https://doi.org/10.17705/1thci.00141 DOI: 10.17705/1thci.00141

Williams, M., & Burden, R. (1997). *Psychology for Language Teachers: A Social Constructivist Approach.* Cambridge University Press, Cambridge

Xu, J., Lio, A., Dhaliwal, H., Andrei, S., Balakrishnan, S., Nagani, U., & Samadder, S. (2021). Psychological interventions of virtual gamification within academic intrinsic motivation: A systematic review. *J Affect Disord*, 293, 444-465. doi: 10.1016/j.jad.2021.06.070.

Yıldırım, İ. (2016). Oyunlaştırma temelli "öğretim ilke ve yöntemleri" dersi öğretim programının geliştirilmesi, uygulanması ve değerlendirilmesi. Gaziantep.

Yıldırım, İ., & Demir, S. (2014). Oyunlaştırma ve eğitim. 11(1), 655-6670.

Yıldırım, İ., & Demir, S. (2015). Oyunlaştırma temelli "öğretim ilke ve yöntemleri" dersi öğretim programı hakkında öğrenci görüşleri. 3rd International Congress on Curriculum and Instruction (ICCI-¬2015).

Yılmaz, E., & Çağıltay, K. (2004). Elektronik oyunlar ve Türkiye. Ulusal Bilişim Kurultayı.

Ye, S., Ng, T., Yim, K., & Wang, J. (2015). Validation of the curiosity and exploration inventory-ii (cei-ii) among Chinese university students in Hong Kong. *J Pers Assess*, 97(4), 403-410.

You, C. J., & Chan, L. (2015). The dynamics of L2 imagery in future motivational self guides. In Z. Dörnyei, P. McIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning*, 397-418. Bristol, UK: Multilingual Matters.

Zhang, L. (2021). Gamification design in self-paced online courses for adult learners: a mixed methods study. PhD thesis, UNSPECIFIED.

Zichermann, G., & Cunningham, C. (2011). *Gamification by design: implementing game mechanics in web and mobile apps*. Sebastopol: O"Reilly Media.

Zichermann, G., & Linder, J. (2010). *Game-based marketing: inspire customer loyalty through rewards, challenges, and contests.* Hoboken,NJ: Wiley.

Zijlstra, W. P. (2009). *Outlier detection in test and questionnaire data for attribute measurement.* Ridderkerk: Ridderprint

Zimbardo, P. G., & Boyd, J. N. (1999). Putting time in perspective: a valid, reliable, individual-differences metric. *Journal of Personality and Social Psychology*, 77(6), 1271-1288.

APPENDIX-A: Informed Consent Form 1

ARAŞTIRMA GÖNÜLLÜ KATILIM FORMU

Bu çalışma Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü doktora öğrencisi Mehmet ABİ tarafından Prof. Dr. İsmail Hakkı ERTEN danışmanlığında yürütülmektedir. Doktora tezi kapsamında yürütülen bu çalışmada oyunlaştırma alanında geliştirilen modellerden birisi olan Octalysis modelinin kuramsal altyapısında yer alan ve motivasyonu etkilediği belirtilen faktörlerin yabancı dil eğitimi alanında geçerli olup olmadıklarının belirlenmesi amaçlanmaktadır. Bu amaçla sizden çevirimiçi ortamda ya da tercihinize göre basılı olarak size sunulacak ölçekteki maddeler ile ilgili olarak kendi düşüncelerinizi en iyi yansıttığını düşündüğünüz seçeneği işaretlemenizi rica ediyoruz. Bu işlem yaklaşık 25-30 dakika sürecektir.

Çalışma için Hacettepe Üniversitesi etik komisyonu ve Eğitim Bilimleri Enstitüsünden gerekli izinler alınmıştır. Çalışmaya katılım gönüllülük esasına dayanmaktadır. Çalışmaya katılıp katılmamayı seçme hakkınız bulunmaktadır. Ayrıca bu çalışmaya katılabilmek için en az 18 yaşında olmanız gerekmektedir. Çalışmaya katılmak istemiyorsanız ya da 18 yaşından küçükseniz lütfen bu formu iade ediniz ya da boş bırakınız.

Veri toplama aracında kişisel risk oluşturacak ya da size rahatsızlık verecek sorular bulunmamaktadır. Ancak, katılım esnasında herhangi bir sebepten dolayı kendinizi rahatsız hissederseniz ya da herhangi bir sebepten dolayı çalışmada yer almak istemezseniz, istediğiniz anda vazgeçmekte ve çalışmadan ayrılmakta serbest olduğunuzu ve bu durumun size hiçbir sorumluluk getirmeyeceğini unutmayınız. Gönüllü katılım formunu imzaladıktan sonra çalışmadan çıkmak isterseniz bunu araştırmacıya söylemeniz yeterli olacaktır. Rahatsızlığınızın olması halinde bu rahatsızlığın giderilmesi konusunda görevli kişiler size yadım etmeye hazırdırlar, lütfen yardım istemekten çekinmeyiniz.

Araştırmadan elde edilen veriler ve kimlik bilgileriniz yalnızca bilimsel amaçlarla kullanılacak ve kimseyle paylaşılmayacaktır. Çalışma ile ilgili olarak aklınıza gelebilecek sorularınız için görevli kişilerle irtibata geçmekten çekinmeyiniz. Ayrıca çalışma bittikten sonra aklınıza gelen sorular olması ya da çalışmanın sonuçları hakkında bilgi almak istemeniz durumunda lütfen aşağıda verilen iletişim adreslerinden irtibata geçmekten çekinmeyiniz. Ayırdığınız vakit için teşekkür ederiz.

Tarih:

Sorumlu Araştırmacının		Yardımcı Araştırmacının	
Adı Soyadı	: Prof. Dr. İsmail Hakkı ERTEN	Adı Soyadı	: Mehmet ABİ
Adres	: Hacettepe Üniversitesi	Adres	: Muğla Sıtkı Koçman Üniversitesi
Tel No.	:	Tel No.	:
e-posta	:	e-posta	:
İmza	:	İmza	:

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda bırakıp çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum.

Katılımcı: Adı, soyadı :

Adres :

APPENDIX-B: Informed Consent Form 2

UYGULAMA GÖNÜLLÜ KATILIM FORMU

Bu çalışma Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü doktora öğrencisi Mehmet ABİ tarafından Prof. Dr. İsmail Hakkı ERTEN danışmanlığında yürütülmektedir. Doktora tezi kapsamında yürütülen bu çalışmada oyunlaştırma alanında geliştirilen modellerden birisi olan Octalysis modelinin kuramsal altyapısında yer alan ve motivasyonu etkilediği belirtilen faktörlerin yabancı dil eğitimi alanında geçerli olup olmadıklarının belirlenmesi amaçlanmaktadır. Bu doğrultuda sekiz haftalık oyunlaştırma uygulaması yapılması planlanmaktadır. Sizden oyunlaştırma uygulaması yapılacak sınıfta yer alarak çalışmaya destek vermenizi rica ediyoruz.

Çalışma için Hacettepe Üniversitesi etik komisyonu ve Eğitim Bilimleri Enstitüsünden gerekli izinler alınmıştır. Çalışmaya katılım gönüllülük esasına dayanmaktadır. Çalışmaya katılıp katılmamayı seçme hakkınız bulunmaktadır. Ayrıca bu çalışmaya katılabilmek için en az 18 yaşında olmanız gerekmektedir. Çalışmaya katılmak istemiyorsanız ya da 18 yaşından küçükseniz lütfen bu formu iade ediniz ya da boş bırakınız.

Uygulama esnasında kişisel risk oluşturacak ya da size rahatsızlık verecek unsurlar bulunmamaktadır. Ancak, katılım esnasında herhangi bir sebepten dolayı kendinizi rahatsız hissederseniz ya da herhangi bir sebepten dolayı çalışmada yer almak istemezseniz, istediğiniz anda vazgeçmekte ve çalışmadan ayrılmakta serbest olduğunuzu ve bu durumun size hiçbir sorumluluk getirmeyeceğini unutmayınız. Uygulama gönüllü katılım formunu imzaladıktan sonra çalışmadan çıkmak isterseniz bunu araştırmacıya söylemeniz yeterli olacaktır. Rahatsızlığınızın olması halinde bu rahatsızlığın giderilmesi konusunda görevli kişiler size yadım etmeye hazırdırlar, lütfen yardım istemekten çekinmeyiniz.

Araştırmadan elde edilen veriler ve kimlik bilgileriniz yalnızca bilimsel amaçlarla kullanılacak ve kimseyle paylaşılmayacaktır. Çalışma ile ilgili olarak aklınıza gelebilecek sorularınız için görevli kişilerle irtibata geçmekten çekinmeyiniz. Ayrıca çalışma bittikten sonra aklınıza gelen sorular olması ya da çalışmanın sonuçları hakkında bilgi almak istemeniz durumunda lütfen aşağıda verilen iletişim adreslerinden irtibata geçmekten çekinmeyiniz. Ayırdığınız vakit için teşekkür ederim.

Tarih:

Sorumlu Araştırmacının		Yardımcı Araştırmacının	
Adı Soyadı	: Prof. Dr. İsmail Hakkı ERTEN	Adı Soyadı	: Mehmet ABİ
Adres	: Hacettepe Üniversitesi	Adres	: Muğla Sıtkı Koçman Üniversitesi
Tel No.	:	Tel No.	:
e-posta	:	e-posta	:
İmza	:	İmza	:

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda bırakıp çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum.

Katılımcı: Adı, soyadı :
Adres :
Tel :
İmza :

APPENDIX-C: Informed Consent Form 3

ARAŞTIRMA GÖNÜLLÜ KATILIM FORMU

Bu çalışma Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü doktora öğrencisi Mehmet ABİ tarafından Prof. Dr. İsmail Hakkı ERTEN danışmanlığında yürütülmektedir. Doktora tezi kapsamında yürütülen bu çalışmada oyunlaştırma alanında geliştirilen modellerden birisi olan Octalysis modelinin kuramsal altyapısında yer alan ve motivasyonu etkilediği belirtilen faktörlerin yabancı dil eğitimi alanında geçerli olup olmadıklarının belirlenmesi amaçlanmaktadır. Bu amaçla sizden çevirimiçi ortamda ya da tercihinize göre basılı olarak size sunulacak ölçekteki maddeler ile ilgili olarak kendi düşüncelerinizi en iyi yansıttığını düşündüğünüz seçeneği işaretlemenizi rica ediyoruz. Bu işlem yaklaşık 25-30 dakika sürecektir.

Çalışma için Hacettepe Üniversitesi etik komisyonu ve Eğitim Bilimleri Enstitüsünden gerekli izinler alınmıştır. Çalışmaya katılım gönüllülük esasına dayanmaktadır. Çalışmaya katılıp katılmamayı seçme hakkınız bulunmaktadır. Ayrıca bu çalışmaya katılabilmek için en az 18 yaşında olmanız gerekmektedir. Çalışmaya katılmak istemiyorsanız ya da 18 yaşından küçükseniz lütfen bu formu iade ediniz ya da boş bırakınız.

Veri toplama aracında kişisel risk oluşturacak ya da size rahatsızlık verecek sorular bulunmamaktadır. Ancak, katılım esnasında herhangi bir sebepten dolayı kendinizi rahatsız hissederseniz ya da herhangi bir sebepten dolayı çalışmada yer almak istemezseniz, istediğiniz anda vazgeçmekte ve çalışmadan ayrılmakta serbest olduğunuzu ve bu durumun size hiçbir sorumluluk getirmeyeceğini unutmayınız. Gönüllü katılım formunu imzaladıktan sonra çalışmadan çıkmak isterseniz bunu araştırmacıya söylemeniz yeterli olacaktır. Rahatsızlığınızın olması halinde bu rahatsızlığın giderilmesi konusunda görevli kişiler size yadım etmeye hazırdırlar, lütfen yardım istemekten çekinmeyiniz.

Araştırmadan elde edilen veriler ve kimlik bilgileriniz yalnızca bilimsel amaçlarla kullanılacak ve kimseyle paylaşılmayacaktır. Çalışma ile ilgili olarak aklınıza gelebilecek sorularınız için görevli kişilerle irtibata geçmekten çekinmeyiniz. Ayrıca çalışma bittikten sonra aklınıza gelen sorular olması ya da çalışmanın sonuçları hakkında bilgi almak istemeniz durumunda lütfen aşağıda verilen iletişim adreslerinden irtibata geçmekten çekinmeyiniz. Ayırdığınız vakit için teşekkür ederiz.

Tarih:

Sorumlu Araştırmacının		Yardımcı Araştırmacının	
Adı Soyadı	: Prof. Dr. İsmail Hakkı ERTEN	Adı Soyadı	: Mehmet ABİ
Adres	: Hacettepe Üniversitesi	Adres	: Muğla Sıtkı Koçman Üniversitesi
Tel No.	:	Tel No.	:
e-posta	:	e-posta	:
İmza	:	İmza	:

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda bırakıp çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum.

Katılımcı: Adı, soyadı :

Adres :

Tel :

APPENDIX-D: Informed Consent Form 4

GÖRÜŞME GÖNÜLLÜ KATILIM FORMU

Bu çalışma Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü doktora öğrencisi Mehmet ABİ tarafından Prof. Dr. İsmail Hakkı ERTEN danışmanlığında yürütülmektedir. Doktora tezi kapsamında yürütülen bu çalışmada oyunlaştırma alanında geliştirilen modellerden birisi olan Octalysis modelinin kuramsal altyapısında yer alan ve motivasyonu etkilediği belirtilen faktörlerin yabancı dil eğitimi alanında geçerli olup olmadıklarının belirlenmesi amaçlanmaktadır. Bu doğrultuda 2018 yılının Eylül, Ekim ve Kasım aylarında sekiz haftalık oyunlaştırma uygulaması planlanmaktadır. Bu amaçla yürütülecek oyunlaştırma uygulaması hakkındaki görüşlerinizi ortaya çıkarmak adına sizlerden haftalık görüşmelere katılarak yarı yapılandırılmış görüşme sorularına yanıt vermenizi rica ediyoruz. Sekiz hafta boyunca cuma günleri Muğla Sıtkı Koçman Üniversitesi Yabancı Diller Yüksekokulunda araştırmacı tarafından yürütülecek görüşmelerde her bir oturumunun 40-60 dk. süreceği tahmin edilmektedir.

Çalışma için Hacettepe Üniversitesi etik komisyonu ve Eğitim Bilimleri Enstitüsünden gerekli izinler alınmıştır. Çalışmaya katılım gönüllülük esasına dayanmaktadır. Çalışmaya katılıp katılmamayı seçme hakkınız bulunmaktadır. Ayrıca bu çalışmaya katılabilmek için en az 18 yaşında olmanız gerekmektedir. Çalışmaya katılmak istemiyorsanız ya da 18 yaşından küçükseniz lütfen bu formu iade ediniz ya da boş bırakınız.

Görüşmeler esnasında ses kaydı alınacaktır. Ancak alınan ses kayıtları yazıya döküldükten sonra silinecektir. Görüşmelerde veri toplama süresince kişisel risk oluşturacak ya da size rahatsızlık verecek sorular bulunmamaktadır. Ancak, katılım esnasında herhangi bir sebepten dolayı kendinizi rahatsız hissederseniz ya da herhangi bir sebepten dolayı çalışmada yer almak istemezseniz, istediğiniz anda vazgeçmekte ve çalışmadan ayrılmakta serbest olduğunuzu ve bu durumun size hiçbir sorumluluk getirmeyeceğini unutmayınız. Görüşme gönüllü katılım formunu imzaladıktan sonra çalışmadan çıkmak isterseniz bunu araştırmacıya söylemeniz yeterli olacaktır. Rahatsızlığınızın olması halinde bu rahatsızlığın giderilmesi konusunda görevli kişiler size yadım etmeye hazırdırlar, lütfen yardım istemekten çekinmeyiniz.

Araştırmadan elde edilen veriler ve kimlik bilgileriniz yalnızca bilimsel amaçlarla kullanılacak ve kimseyle paylaşılmayacaktır. Çalışma ile ilgili olarak aklınıza gelebilecek sorularınız için görevli kişilerle irtibata geçmekten çekinmeyiniz. Ayrıca çalışma bittikten sonra aklınıza gelen sorular olması ya da çalışmanın sonuçları hakkında bilgi almak istemeniz durumunda lütfen aşağıda verilen iletişim adreslerinden irtibata geçmekten çekinmeyiniz. Ayırdığınız vakit için teşekkür ederiz.

Tarih:

Sorumlu Araştırmacının		Yardımcı Araştırmacının	
Adı Soyadı	: Prof. Dr. İsmail Hakkı ERTEN	Adı Soyadı	: Mehmet ABİ
Adres	: Hacettepe Üniversitesi	Adres	: Muğla Sıtkı Koçman Üniversitesi
Tel No.	:	Tel No.	:
e-posta	:	e-posta	:
İmza	:	İmza	:

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda bırakıp çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum.

Katılımcı: Adı, soyadı :

Adres :

Tel :

APPENDIX-E: Achievement Goal Questionnaire (Tur)

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, *100* (3), 613-628.

Ölçek: 1-5 (Kesinlikle katılmıyorum – Kesinlikle katılıyorum)

Amacım sınıfta öğretilen konulara tam olarak hâkim olmaktır.

Diğer öğrencilerle kıyaslandığında daha çok başarmak için gayret ederim.

Amacım mümkün olduğunca çok şey öğrenmektir.

Amacım diğer öğrencilere kıyasla iyi performans göstermektir.

Amacım yapabileceğimden daha azını öğrenmekten kaçınmaktır.

Amacım diğer öğrencilere kıyasla zayıf performans göstermekten kaçınmaktır.

Konuyu tam olarak anlamaya gayret gösteririm.

Amacım diğer öğrencilerden daha iyi performans göstermektir.

Amacım öğrenilmesi gerekenden daha az öğrenmekten kaçınmaktır.

Diğer öğrencilerden daha kötü performans göstermekten kaçınmaya gayret ederim.

Ders konularını yarım yamalak anlamaktan kaçınmaya gayret ederim.

Amacım diğer öğrencilerden daha başarısız olmaktan kaçınmaktır.

APPENDIX-F: Achievement Goal Questionnaire (Eng)

Achievement Goal Questionnaire-Revised (AGQ-R)

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. Journal of Educational Psychology, 100 (3), 613-628.

SCALE: 1-5 (Strongly Disagree to Strongly Agree)

My aim is completely master the material presented in this class.

I am striving to do well compared to other students.

My goal is to learn as much as possible.

My aim is to perform well relative to other students.

My aim is to avoid learning less than I possibly could.

My goal is to avoid performing poorly compared to others.

I am striving to understand the content as thoroughly as possible.

My goal is to perform better than the other students.

My goal is to avoid learning less than it is possible to learn.

I am striving to avoid performing worse than others.

I am striving to avoid an incomplete understanding of the course material.

My aim is to avoid doing worse than other students.

APPENDIX-G: Basic Psychological Needs Satisfaction and Frustration (Tur)

TEMEL PSİKOLOJİK İHTİYAÇLARIN TATMİNİ VE ÖFKE ÖLÇEĞİ – GENEL ÖLÇEK

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Duriez, B. Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., & Van der Kaap-Deeder, J., Verstuyf, J (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. Motivation and Emotion, 39, 216-236.

Ölçek 1-5 (Hiç doğru değil – Tamamen doğru)

1. Üstlendiğim işlerde seçim şansım olduğunu ve özgür olduğumu hissediyorum.
2. Yaptığım çoğu şeyi yapmak zorundaymışım gibi hissediyorum.
3. Önemsediğim insanların da beni önemsediklerini hissediyorum.
4. İçinde olmak istediğim gruptan dışlandığımı hissediyorum.
5. İşleri iyi yapabileceğim konusunda kendimden eminim.
6. İşleri iyi yapıp yapamayacağım konusunda ciddi şüphelerim var.
7. Verdiğim kararların gerçekten istediğim şeyleri yansıttığını hissediyorum.
8. Yapmayı seçmeyeceğim birçok şeyi yapmak zorunda bırakıldığımı hissediyorum.
9. Benim önemsediğim ve beni önemseyen insanlarla bağlantıda olduğumu hissediyorum.
10. Benim için önemli olan insanların bana karşı soğuk ve mesafeli olduklarını hissediyorum.
11. Yaptığım şeyde yetenekli olduğumu hissediyorum.
12. Sergilediğim performansımın çoğunun bende hayal kırıklığı yarattığını hissediyorum.
13. Yaptığım seçimlerin beni ifade ettiğini hissediyorum.
14. Birçok şeyi yapmada baskı altında olduğumu hissediyorum.
15. Benim için önemli olan insanlara karşı kendimi onlara yakın ve onlarla bağlantılı hissediyorum.
16. Beraber vakit geçirdiğim insanların beni sevmediklerini zannediyorum.
17. Amaçlarımı gerçekleştirebileceğim konusunda kendimden eminim.
18. Yeteneklerime güvenmediğimi hissediyorum.
19. Gerçekten ilgim olan şeyleri yapmakta olduğumu hissediyorum.
20. Günlük yaptığım işleri sanki bir zorunluluklar zinciri gibi olduğunu hissediyorum.
21. Beraber vakit geçirdiğim insanlara karşı samimi duygular yaşıyorum.
22. Sahip olduğum ilişkilerin yüzeysel olduğunu hissediyorum.
23. Zor görevleri başarı ile tamamlayabileceğimi hissediyorum.
24. Yaptığım hatalara yüzünden başarısız olmuş hissediyorum.

APPENDIX-H: Basic Psychological Needs Satisfaction and Frustration (Eng)

General Measure – English Version
Below, we are going to ask about your actual experiences of certain feelings in your life.
Please read each of the following items carefully. You can choose from 1 to 5 to indicate the degree to which the statement is true for you at this point in your life.
1 2 3 4 5
Not True at all Completely
True
1. I feel a sense of choice and freedom in the things I undertake 1 2 3 4 5
2. Most of the things I do feel like “I have to” 1 2 3 4 5
3. I feel that the people I care about also care about me 1 2 3 4 5
4. I feel excluded from the group I want to belong to 1 2 3 4 5
5. I feel confident that I can do things well 1 2 3 4 5
6. I have serious doubts about whether I can do things well 1 2 3 4 5
7. I feel that my decisions reflect what I really want 1 2 3 4 5
8. I feel forced to do many things I wouldn’t choose to do 1 2 3 4 5
9. I feel connected with people who care for me, and for whom I care 1 2 3 4 5
10. I feel that people who are important to me are cold and distant towards me
1 2 3 4 5
11. I feel capable at what I do 1 2 3 4 5
12. I feel disappointed with many of my performance 1 2 3 4 5
13. I feel my choices express who I really am 1 2 3 4 5
14. I feel pressured to do too many things 1 2 3 4 5
15. I feel close and connected with other people who are important to me. 1 2 3 4 5
16. I have the impression that people I spend time with dislike me 1 2 3 4 5
17. I feel competent to achieve my goals
1 2 3 4 5
18. I feel insecure about my abilities 1 2 3 4 5
19. I feel I have been doing what really interests me 1 2 3 4 5
20. My daily activities feel like a chain of obligations 1 2 3 4 5
21. I experience a warm feeling with the people I spend time with 1 2 3 4 5
22. I feel the relationships I have are just superficial 1 2 3 4 5
23. I feel I can successfully complete difficult tasks 1 2 3 4 5
24. I feel like a failure because of the mistakes I make 1 2 3 4 5
Scoring
Autonomy satisfaction: items 1, 7, 13, 19
Autonomy frustration items: 2, 8, 14, 20
Relatedness satisfaction: items 3, 9, 15, 21
Relatedness frustration items 4, 10, 16, 22
Competence satisfaction: items 5, 11, 17, 23
Competence frustration items 6, 12, 18, 24

APPENDIX-I: Curiosity and Exploration Inventory II (Tur)

Merak ve Keşfetme Envanteri

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The Curiosity and Exploration Inventory-II. Development, factor structure, and psychometrics. Journal of Research in Personality, 43, 987-998.

Ölçek 1-5 (Çok az ya da Hiç – Çok çok fazla)

1. Yeni bir şeylerle karşılaştığım durumlarda olabildiğince çok bilgi edinmenin yollarını ararım.
2. Günlük yaşamın belirsizliklerini seven türde birisiyimdir.
3. Karmaşık ya da zorlu şeyleri yapmada çok iyiyimdir.
4. Gittiğim her yerde yeni bir şeyler ya da deneyimler ararım.
5. Zorlu şeyleri birer gelişme ve öğrenme fırsatı olarak görürüm.
6. Birazcık korkutucu olan şeyleri yapmayı severim.
7. Her zaman kedimle ilgili düşüncelerimi ya da dünya görüşümü zorlayıcı deneyimleri ararım.
8. Heyecan verici ama aynı zamanda tahmin edilemez olan işleri tercih ederim.
9. Sık sık beni zorlayan ve kişisel gelişimime katkısı olacak fırsatları ararım.
10. Tanımadığım kişi, olay ve yerleri kucaklayan tarzda birisiyimdir.

APPENDIX-J: Curiosity and Exploration Inventory II (Eng)

Curiosity and Exploration Inventory (CEI-II)
Instructions: Rate the statements below for how accurately they reflect the way you
generally feel and behave. Do not rate what you think you should do, or wish you do, or
things you no longer do. Please be as honest as possible.

Very Slightly or Not At All
A Little
Moderately
Quite a Bit
Extremely

1. I actively seek as much information as I can in new situations. 1 2 3 4 5
2. I am the type of person who really enjoys the uncertainty of everyday life. 1 2 3 4 5
3. I am at my best when doing something that is complex or challenging. 1 2 3 4 5
4. Everywhere I go, I am out looking for new things or experiences. 1 2 3 4 5
5. I view challenging situations as an opportunity to grow and learn. 1 2 3 4 5
6. I like to do things that are a little frightening. 1 2 3 4 5
7. I am always looking for experiences that challenge how I think about myself and the world. 1 2 3 4 5
8. I prefer jobs that are excitingly unpredictable. 1 2 3 4 5
9. I frequently seek out opportunities to challenge myself and grow as a person. 1 2 3 4 5
10. I am the kind of person who embraces unfamiliar people, events, and places. 1 2 3 4 5

Stretching: 1,3,5,7 / Embracing: 2,4,6,8.10.

APPENDIX-K: DMC Disposition Questionnaire (Tur)

YÖNLENDİRİLMİŞ MOTİVASYON AKIMLARI ANKETİ

Uzun Vadeli Motivasyonu Anlamak

Araştırmacılara göre, insanlar bir seferde haftalar ve hatta aylar boyunca kendilerini bir işe kaptırmakta ve ÇOK YOĞUN bir şekilde o işe odaklanabilmektedirler. Bu tür durumlarda kişilerin kendilerini tamamen ilgilendikleri işe verdikleri görülmektedir ve genelde içinde bulundukları durumu şu sözlerle ifade ettikleri gözlenmektedir:

- Gece gündüz bu projeyi düşünüyorum. Sanki tüm hayatım bu işten ibaret.
- Bu kadar uzun süre, bir şeye odaklanabilmek müthiş bir şey. O kadar güzel ki her şey çok kolay geliyor.
- Hiç bu kadar başarılı olabileceğimi düşünmemiştim.
- Arkadaşlarım kesinlikle bende değişik bir şey olduğunu fark ediyorlar. Benim başka bir şeye kendimi bu kadar verdiğimi görmediklerini söylüyorlar.
- Keşke kendimi yaptığım diğer işlere de bu kadar verebilseydim. O zaman hedeflerime ulaşmak daha kolay olurdu.

İngilizce öğrenme sürecinizde, çalışırken yukarıda bahsedilen gibi motivasyon yoğunluğunu ne sıklıkla yaşarsınız?

◯ Bu tür bir motivasyonu HİÇ yaşamadım (eğer cevabınız Hiç ise lütfen .. soruya geçiniz.)

◯ BIR KEREsinde böyle bir motivasyon yaşadım ama YUKARIDAKI KADAR YOĞUN DEĞILDI.

◯ BIRKAÇ KERE böyle bir motivasyon yaşadım ama YUKARIDAKI KADAR YOĞUN DEĞILDI.

◯ Bu tür bir motivasyonu BIR KERE yukarıda bahsedilene BENZER YOĞUNLUKTA yaşadım.

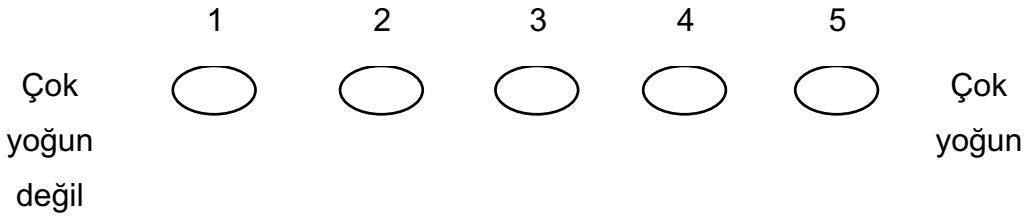
◯ Bu tür bir motivasyonu BIRKAÇ KERE yukarıda bahsedilene BENZER YOĞUNLUKTA yaşadım.

Yaşamış olduğunuz bu yoğunluktaki motivasyonunuz hakkında

Eğer birkaç kere bu yoğunlukta bir motivasyon deneyimi yaşadıysanız, lütfen en iyi hatırladığınızı seçiniz. Bu deneyim ne kadar sürdü?

☐	1 aydan az
☐	1-2 ay
☐	2-4 ay
☐	4-6 ay
☐	6 aydan fazla

1. Lütfen aşağıdaki ölçek üzerinde bu süre içerisinde motivasyonunuzun ne kadar yoğun olduğunu belirtin.



2. Yaşadığınız bu yoğun motivasyon hakkında

	Kesinlikle Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
1. Geriye dönüp baktığımda bu süreçle ilgili güzel anılarım var.					
2. Bu süre içerisinde her zamankinden daha verimli bir şekilde çalışabilmiştim.					
3. Yapabildiğim şeye ben bile şaşırmıştım.					
4. Çoğu zaman devam etmek çok zor gelmişti.					
5. Bu deneyim benim istediğimden de çok başarılı olmama yardımcı oldu.					
6. Sanırım bu süreç içerisinde bana özel bir şey oldu – çok müthiş bir zamandı.					
7. O zaman proje benim hayatımın merkezi haline gelmişti.					
8. Çevremdeki insanlar özel bir şeyler yaşadığımı fark etmişlerdi.					
9. O zaman çok çalışıyormuşum gibi gelmemişti. Sanki bir akıntıda gibiydim. Kendimi kaptırmıştım.					
10. Sürekli hedefimi düşündüğümü hatırlıyorum.					
11. Sık sık hedefime ulaştığımı hayal ediyordum.					
12. Gerçekten çok eğlenceli bir deneyimdi.					

APPENDIX-L: DMC Disposition Questionnaire (Eng)

Understanding long-term motivation

Hello!

My name is Christine Muir and I am a researcher from the University of Nottingham, England. I am carrying out research into long-term motivation and would be very grateful if you could help me by answering a few questions.

It will not take more than 7-10 minutes, and your answers will be invaluable in helping me understand why and how people engage in long-term projects. I will use this information to help us learn more about how we can better support students to succeed in their studies.

Thank you so much for all your help!

Best wishes,
Christine

Understanding long-term motivation

*Required

About your intense motivational project

*

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
When looking back now, I have very good memories of this time	◯	◯	◯	◯	◯
During this time I was able to work more productively than I usually can	◯	◯	◯	◯	◯
I surprised myself with how much I was able to do	◯	◯	◯	◯	◯
Many times it felt like a real struggle to keep going	◯	◯	◯	◯	◯
This experience helped me to achieve all I had wanted to and more	◯	◯	◯	◯	◯
I think something special happened to me during this experience - it was an amazing time	◯	◯	◯	◯	◯

Would you like to experience this type of intense motivation again? *

◯ Yes

◯ No

Would you mind telling us briefly why? Thank you!

« Back Continue »

44% completed

Understanding long-term motivation

*Required

A little more about your intensive project experience...

*

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
At the time, this project became a central part of my life	◯	◯	◯	◯	◯
The people around me could see that I was experiencing something special	◯	◯	◯	◯	◯
It didn't feel like hard work at the time - I was just caught up in the flow!	◯	◯	◯	◯	◯
I remember thinking about my goal all the time	◯	◯	◯	◯	◯
I often imagined myself achieving my final goal	◯	◯	◯	◯	◯
It was a really enjoyable experience	◯	◯	◯	◯	◯

« Back Continue »

55% completed

APPENDIX-M: Susceptibility to Persuasion (Tur)

İkna Edilmeye Yatkınlık Ölçeği

Kaptein, M., Ruyter, B., Markopoulos, P., & Aarts, E. (2012). Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking. *ACM Transactions on Interactive Intelligent Systems*, 1025.

Ölçek 1-7 (Tamamen katılmıyorum – Tamamen katılıyorum)

1. Bence az bulunan (nadir) ürünler her yerde bulunan ürünlerden daha değerlidir.
2. Favori dükkanım kapancak üzereyse, son şansım olduğu için mutlaka uğramaya çalışırım.
3. Bir ürün alabilen en son kişi bensem kendimi iyi hissederim.
4. Favori şampuanım satıcının stoklarında tükenmek üzereyse iki şişe alırım.
5. Elde edilmesi zor ürünlerin kendilerine has özel değerleri vardır.

APPENDIX-N: Susceptibility to Persuasion (Eng)

Susceptibility to Persuasion

Kaptein, M., Ruyter, B., Markopoulos, P., & Aarts, E. (2012). Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking. *ACM Transactions on Interactive Intelligent Systems*, 1025.

Scale 1-7 (Strongly Disagree – Strongly Agree)

Scarcity

Scarce 6∗ I believe rare products (scarce) are more valuable than mass products.

Scarce 7∗ When my favorite shop is about to close, I would visit it since it is my last chance.

Scarce 8∗ I would feel good if I was the last person to be able to buy something.

Scarce 9∗ When my favorite shampoo is almost out of stock I buy two bottles.

Scarce 10∗ Products that are hard to get represent a special value.

APPENDIX-O: Psychological Ownership Questionnaire (Tur)

Psikolojik Sahiplik Anketi

Astaryan, V. S., Slevitch, L., Larzelere, R., Morosan, C., & Kwun, D. J. (2013). ffects of Psychological Ownership on Students'Commitment and Satisfaction. *ournal of Hospitality & Tourism Education, 25*(4), 169-179.

Aşağıdaki ifadeler sizin akademik programa katılmanızla ilgilidir. Lütfen aşağıdaki ölçeği kullanarak sizce uygun seçeneği işaretleyiniz.

Deneyimlerinize dayanarak, aşağıdakilerin her birini akademik programınızda ne sıklıkla yaptınız.

Ölçek 1-5 (Asla – Çok sık)

1. Sınıf projelerinde diğer öğrencilerle birlikte çalıştınız.
2. Görevleri (ör. ev ödevi, projeler, sınıf tartışmaları vb.) tamamlarken fikir vererek katkıda bulundunuz.
3. Sınıf dışında diğer kişilerle (öğrenciler, aile üyeleri, beraber çalıştığınız kişiler vb) programda öğrendiğiniz şeyleri tartıştınız.
4. Okul dışı aktivitelere katıldınız (organizasyonlar, yayınlar, öğrenci klüpleri vb.).
5. Sınıfta sunum yaptınız.

Aşağıdaki sorular sizin içinde bulunduğunuz akademik programa karşı hisleriniz ile ilgilidir. Lütfen aşağıdaki ölçeği kullanarak ifadelere ne derece katılıp katılmadığınızı belirtiniz.

Ölçek 1-5 (Kesinlikle katılmıyorum – Kesinlikle katılıyorum)

1. Bu programa ait olduğumu hissediyorum.
2. Bu programda kendimi "evimdeymiş" gibi hissediyorum.
3. Bu programda bir yabancı gibi hissediyorum.
4. Programa bağlı olduğumu hissediyorum.
5. Bu programda aldığım eğitimin kalitesi üzerinde etkim var.
6. Bu programda aldığım eğitimin kontrolünün bende olduğunu hissediyorum.
7. Bu program içerisinde çalışmalarımı düzenleyebilirim.
8. Bu programda nasıl öğrenmem üzerinde çok etkim var.
9. Bu programdan diploma alacak olmaktan gurur duyuyorum.
10. Başkalarına bu programda olduğumu söylemekten gurur duyuyorum.
11. Bu programın eski mezunlarından birisi olmaktan gurur duyuyorum.
12. Bu programı parlak öğrencilere gurur duyarak önerebilirim.
13. Bu programa katılma konusunda vermiş olduğum karardan son derece memnunum.
14. Bu programa kaydolma konusunda verdiğim karar akıllıcaydı.
15. Bu programa kaydolduğum için mutluyum.
16. Eğer üniversiteye tekrar başlayabilseydim gene bu programı seçerdim.
17. Bu programdan mezun olmak benim için önemli.
18. Bu programın geleceği ile gerçekten ilgileniyorum.
19. Mezuniyetimden sonra program ile iş birliği yapmaya niyetliyim.
20. Mezuniyetimden sonra kuruma finansal katkı sağlama niyetindeyim (burs vermek, kütüphaneye destek olmak, sosyal projeler vb.).

APPENDIX-P: Psychological Ownership Questionnaire (Eng)

Psychological Ownership Survey Questions

In your experience, how often have you done each of the following in your academic program?

1 = Never 2 3 4 5= Very Often

1. Worked with other students on class projects.
2. Contributed ideas when completing assignments (i.e., homework, projects, and class discussions).
3. Discussed what you have learned in the program with others outside the class (students, family members, coworkers, etc.)
4. Participated in extra-curricular activities (organizations, publications, student honor societies, etc.
5. Made a class presentation

The following questions are about your feelings toward your academic program. Please use the scale below to indicate your level of agreement/disagreement with the following statements.

1= Strongly Disagree 2= Disagree 3= Neutral 4=Agree 5=Strongly Agree

1. I feel I belong to this program.
2. I feel "at home" in this program.
3. I feel like a stranger in this program.
4. I feel attached to the program.
5. I have influence over the quality of education I receive in this program.
6. I feel in control over my education in this program.
7. I can modify my plan of study in this program.
8. I have a lot of influence on how my learning occurs in this program.
9. I am proud of getting a degree from this program.
10. I am proud of telling others I am studying in this program.
11. I am proud of becoming a part of the alumni of this program.
12. I will proudly recommend prospective students this program.
13. I am satisfied with my decision to attend this program.
14. My choice to enroll in this program was a wise idea.
15. I am happy that I enrolled in this program.
16. If I could start college/university over, I would choose to attend this program.
17. It is important for me to graduate from *this* program.
18. I really care about the fate of this program.
19. I intend to continue to associate with the program after graduation.
20. I intend to contribute to the institution financially after graduation (e.g., funds for scholarships, library support, special projects, etc.)

APPENDIX-Q: Semi-Structured Interview Questions

1. Bu hafta girdiğiniz derslerde hoşunuza giden uygulamalar var mıydı?
2. Bu hafta girdiğiniz derslerde hoşlanmadığınız uygulamalar var mıydı?
3. Bu hafta girdiğiniz derslerde dikkatinizi çeken bir şey oldu mu?
4. Bu hafta girdiğiniz dersler ile ilgili önerileriniz var mı?

APPENDIX-R: Ethics Committee Approval

T.C.
HACETTEPE ÜNİVERSİTESİ
Rektörlük

Sayı : 35853172/
433 -1654

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi: 19.03.2018 tarih ve 770 sayılı yazınız.

Enstitümüz Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Bilim Dalı doktora programı öğrencilerinden **Mehmet ABİ**'nin **Prof. Dr. İsmail Hakkı ERTEN** danışmanlığında yürüttüğü **"Octalysis ile Dil Öğrenme Ortamlarının Optimize Edilmesinde Oyunlaştırma ve Yönlendirilmiş Motivasyon Akımları"** başlıklı tez çalışması, Üniversitemiz Senatosu Etik Komisyonunun **03 Nisan 2018** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini rica ederim.

Prof. Dr. Rahime M. NOHUTCU
Rektör a.
Rektör Yardımcısı

Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara
Telefon: 0 (312) 305 3001 - 3002 • Faks: 0 (312) 311 9992
E-posta: yazimd@hacettepe.edu.tr • www.hacettepe.edu.tr

Ayrıntılı Bilgi için:
Yazı İşleri Müdürlüğü
0 (312) 305 1008

APPENDIX S: Declaration of Ethical Conduct

I hereby declare that…

- I have prepared this thesis in accordance with the thesis writing guidelines of the Graduate School of Educational Sciences of Hacettepe University;
- all information and documents in the thesis/dissertation have been obtained in accordance with academic regulations;
- all audio visual and written information and results have been presented in compliance with scientific and ethical standards;
- in case of using other people's work, related studies have been cited in accordance with scientific and ethical standards;
- all cited studies have been fully and decently referenced and included in the list of References;
- I did not do any distortion and/or manipulation on the data set,
- and **NO** part of this work was presented as a part of any other thesis study at this or any other university.

(DD) /(MM)/(YY)

Mehmet ABİ

APPENDIX-T: Thesis/Dissertation Originality Report

……/……./……

HACETTEPE UNIVERSITY

Graduate School of Educational Sciences

To The Department of English Language Teaching

Thesis Title: Gamification and Directed Motivational Currents in Optimizing Language Learning Environments through Octalysis

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defence	Similarity Index	Submission ID
22/09 /2021	400	646244	22/10 /2021	12	1654973030

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Name Lastname:	Mehmet Abi
Student No.:	N14243010
Department:	Foreign Language Teaching
Program:	English Language Teaching
Status:	☐ Masters ☒ Ph.D. ☐ Integrated Ph.D.

Signature

ADVISOR APPROVAL

APPROVED
(Dr. Öğr. Üyesi Hatice Ergül)

APPENDIX-U: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım haklan bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan **"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına ilişkin Yönerge"** kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- o Enstitü/ Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. (1)
- o Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren … ay ertelenmiştir. (2)
- o Tezimle ilgili gizlilik kararı verilmiştir. (3)

…… /…… /……

(imza)

Mehmet ABİ

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

(1) Madde 6. 1. Lisansüstü tezle ilgili patent başvurusu yapılması veya patent alma *sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü Üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.*

(2) Madde 6. 2. *Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanın önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.*

(3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü Üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.*
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

** Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.*