LUIZ BERNARDO MARTINS KUMMER

IDENTIFYING PLAYERS' PSYCHOLOGICAL PROFILE IN DIGITAL GAMES BASED ON USAGE DATA

Curitiba, Paraná, Brazil

2021

LUIZ BERNARDO MARTINS KUMMER

IDENTIFYING PLAYERS' PSYCHOLOGICAL PROFILE IN DIGITAL GAMES BASED ON USAGE DATA

Doctoral thesis presented to the Graduate Program in Informatics at the Pontifícia Universidade Católica do Paraná as a partial fulfillment of the requirements for the degree of Doctor in Computer Science.

Pontifícia Universidade Católica do Paraná - PUCPR

Programa de Pós-Graduação em Informática - PPGIa

Supervisor: Emerson Cabrera Paraiso

Co-supervisor: Júlio César Nievola

Curitiba, Paraná, Brazil

2021

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central
Edilene de Oliveira dos Santos CRB-9/1636

Kummer, Luiz Bernardo Martins
K96i Identifying players' psychological profile in digital games based on usage
2021 data / Luiz Bernardo Martins Kummer ; supervisor: Emerson Cabrera Paraiso ; co-supervisor: Júlio César Nievola. -- 2021
320 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2021
Bibliografia: f. 300-317

1. Informática. 2. Jogos para computador. 3. Inteligência artificial. 4. Redes neurais (Computação). 5. Comportamento humano.
I. Paraiso, Emerson Cabrera. II. Nievola, Júlio César.
III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática. IV. Título

. CDD. 20.ed. – 004

Pontifícia Universidade Católica do Paraná
Escola Politécnica
Programa de Pós-Graduação em Informática

ATA DA SESSÃO PÚBLICA DE DEFESA DE TESE DE DOUTORADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA, DA PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ

DEFESA DE TESE N.º 79

À DISTÂNCIA

Em sessão pública realizada à distância, às 08h00 no dia 25 de fevereiro de 2021, ocorreu a defesa da tese de doutorado intitulada **"IDENTIFYING PLAYERS' PSYCHOLOGICAL PROFILE IN DIGITAL GAMES BASED ON USAGE DATA"** elaborada pelo(a) aluno(a) **LUIZ BERNARDO MARTINS KUMMER**, como requisito parcial para a obtenção do título de **Doutor em Informática**, na área de concentração **Ciência da Computação.** Após a apresentação da tese pelo(a) aluno(a) e correspondente arguição, a banca examinadora emitiu o seguinte parecer sobre a tese:

Membros da Banca Examinadora:

Prof. Dr. Emerson Cabrera Paraiso (PUCPR/PPGIa) Orientador
DELIBERAÇÃO: APROVADO ASSINATURA: ____________________

Prof. Dr. Julio Cesar Nievola (PUCPR/PPGIa) Co-orientador
DELIBERAÇÃO: APROVADO ASSINATURA: À DISTÂNCIA

Prof. Dr. Carlos Nascimento Silla Jr. (PUCPR/PPGIa)
DELIBERAÇÃO: APROVADO ASSINATURA: À DISTÂNCIA

Prof. Dr. Esteban Walter Gonzales Clua (Universidade Federal Fluminenese) - UFF
DELIBERAÇÃO: APROVADO ASSINATURA: À DISTÂNCIA

Prof. Dr. Hiroyuki Iida (Japan Advanced Institute of Science and Technology) - JAIST
DELIBERAÇÃO: APROVADO ASSINATURA: À DISTÂNCIA

Portanto, conforme as normas regimentais do PPGIa e da PUCPR, a tese foi considerada:
(x) APROVADA
(aprovação condicionada ao atendimento integral das correções e melhorias recomendadas pela banca examinadora, conforme anexo, dentro do prazo regimental)
( ) REPROVADA

____
Prof. Dr. Emerson Cabrera Paraiso
Coordenador do Programa de Pós-Graduação em Informática

Pontifícia Universidade Católica do Paraná
Escola Politécnica
Programa de Pós-Graduação em Informática

40-2021

DECLARAÇÃO

Declaro para os devidos fins que o aluno **LUIZ BERNARDO MARTINS KUMMER**, defendeu sua tese de Doutorado intitulada **"IDENTIFYING PLAYERS' PSYCHOLOGICAL PROFILE IN DIGITAL GAMES BASED ON USAGE DATA"**, na área de concentração Ciência da Computação, no dia 25 de fevereiro de 2021, no qual foi aprovado.

Declaro ainda que foram feitas todas as alterações solicitadas pela Banca Examinadora, cumprindo todas as normas de formatação definidas pelo Programa.

Por ser verdade, firmo a presente declaração.

Curitiba, 08 de junho de 2021.

Prof. Dr. Emerson Cabrera Paraiso
Coordenador do Programa de Pós-Graduação em Informática
Pontifícia Universidade Católica do Paraná

Rua Imaculada Conceição,1155 Prado Velho CEP 80215-901 Curitiba Paraná Brasil
Fone: (41) 3271-1669 e-mail: secretaria@ppgia.pucpr.br

This work's accomplishment would not be possible without my sources of motivation and inspiration, my wife, Tatiana, and my daughter Rafaela. Thank you for the support in this journey. It was not easy. I love you. I believe that this work is just good preparation for a longer adventure, where we will produce and enjoy good fruits.

Acknowledgements

I start thanking my wife Tatiana and my daughter Rafaela for the unconditional love, care, and support. You are my eternal source of dedication; this work is for you. All companionship during extra work hours on holidays and weekends, waking up early with me all days even when we had terrible sleeping nights, all of it gave me more strength to do my best. I love you all.

I thank my birth family, my father Mauro, who always encouraged me to continue studying, my mom Ana who always wishes me the best, and my sisters Regina e Mila, whom I always keep in mind when putting myself into new challenges. You made me grow up as a good person. I thank all my friends, which always pushed me up and celebrated my achievements with so much energy. I also thank my karate Sensei Silvério, who taught me about discipline, respect, and a sense of effort. Special thanks to my cousin Ittalo which shares the same passion for games as me and helped me many times by verifying this work's extension and applicability. For my friend Cristian, I thank you for all the insightful database lessons that allowed me to perform many kinds of experiments.

All of it would not be possible without the guidance of my advisor Prof. Dr. Emerson Cabrera Paraiso, my co-advisor Prof. Dr. Júlio César Nievola, and my co-advisor Prof. Dr. Hiroyuki Iida in Japan. With less contact, but without less companionship, I thank Prof. Dr. Carlos Nascimento Silla Junior, Prof. Dr. Orlando Alcantara Soares, and Prof. Dra. Andreia Malucelli. This work would not be done without your guidance, thank you very much. I also thank my lab-mate João, that supported me many times by validating my ideas. If I am a good researcher now, it is because of you. In particular, I describe that we always had a special element in our works, the fact that we are not game experts neither psychologists, but based on the pleasure to research, we could contribute on both academia and technology transfer aspects.

Finally, but not with less value, I thank CAPES (Coordenação de Aperfeicoamento de Pessoal de Nivel Superior) for its support in my doctoral degree in Brazil and Japan. I also thank NCSoft for turning the Blade&Soul dataset available for research purposes.

"Com dedicação, tudo é possível."

Abstract

Identifying players' motivations is not a trivial task, demanding metrics to better understand the changes in their behaviors during a game usage lifecycle. Players can change their interest in continuing playing for many reasons, such as the relation between the game content available to them and their personal preferences. Moreover, the variances of their motivations can be interpreted as risk situations or opportunities by game producers. In academia, there are two types of research focusing on players' behavior that are not usually used together: the first one where general psychological models are proposed (textual descriptions of players' behaviors), and the second one where usage tendencies are identified on data (Game Analytics). As the state-of-the-art approaches link only one psychological model to data at a time, we wish to explore the benefits of linking multiple models simultaneously. In this work, we propose to build a novel bridge between the psychological models and Game Analytics. This bridge tries to join numerous psychological models to produce a composition of psychological features (i.e., a profile) through the following problem statement "Is it possible to identify psychological profiles encompassing multiple psychological models on digital games based on usage data?". Besides, 11 research questions are presented to guide the statement's answer. Our findings of over 67 identified psychological models allowed us to propose a method that identifies players' psychological profiles encompassing short, mid, and long-term aspects. The method was applied in an MMORPG (Massively-multiplayer Online Role-playing Game), and the resultant profiles were assessed in a baseline regarding a churn prediction problem. As a result, the proposed arrangement of features outperformed the baseline, showing that it represents players' psychological essences applicable, at least, to the churn problem. The proposed composition also allowed a change of perspective in the churn management from reactive to proactive. The positive method assessment allowed the coining of the term "Psychological Data Enhancement", which references the method conception. Regarding the transfer of technology to the industry, a desktop system named "**P**layer **P**sychological **P**rofile Identification System", or just 3PIS, was developed contemplating all the proposed psychological features processing.

Keywords: Player Behavior, Player Profiling, Player Modeling, Game Usage Lifecycle, Game Analytics, Player Psychological Profile, Systematic Literature Review, Player Psychological Profile Identification System (3PIS).

Resumo

A identificação das motivações dos jogadores não é uma tarefa trivial, sendo realizada com o apoio de métricas para melhor entender as mudanças em seus comportamentos durante um ciclo de vida de uso de um jogo. Os jogadores podem mudar seu interesse em continuar jogando devido a muitas razões, como a relação entre o conteúdo do jogo disponível para eles e suas preferências pessoais. Além disso, tais variações de motivação podem ser interpretadas como situações de risco ou de oportunidade pelos produtores de jogos. No meio acadêmico, existem dois tipos de trabalhos que focam no comportamento dos jogadores que geralmente não são aplicados em conjunto, o primeiro em que modelos psicológicos gerais são propostos (descrições textuais de comportamentos de jogadores), e o segundo onde as tendências de uso são identificadas nos dados (Game Analytics). Como as abordagens do estado da arte vinculam apenas um modelo psicológico aos dados por vez, desejamos explorar os benefícios de se vincular múltiplos modelos ao mesmo tempo. Portanto, neste trabalho, propomos construir uma ponte inovadora entre os modelos psicológicos e as abordagens de Game Analytics através da junção de múltiplos modelos psicológicos para produzir uma composição de características psicológicas (um perfil). Esta proposta é representada pela seguinte declaração de problema "É possível identificar perfis psicológicos abrangendo múltiplos modelos psicológicos em jogos digitais baseados em dados de uso?". Em adição a isso, 11 questões de pesquisa são apresentadas para orientar a resposta da declaração. Nossos achados sobre os 67 modelos psicológicos identificados nos permitiram propor um método que identifica perfís psicológicos de jogadores considerando aspectos de curto, médio e longo prazo. Este método foi aplicado num jogo MMORPG (jogo RPG Massivos de Multijogadores), sendo os perfís resultantes validados numa comparação com o melhor resultado da literatura em relação a uma predição de abandono de jogadores. Como resultado, a composição proposta obteve melhores resultados em comparação ao estado da arte, mostrando que esta composição de características possui essências psicológicas, aplicáveis ao menos, ao problema de abandono de jogadores. A mesma composição possibilitou uma mudança no gerenciamento do abandono de jogadores de uma perspectiva reativa para uma pró-ativa. Dada a validação favorável do método, foi possível cunhar o termo "Psychological Data Enhancement", que refere-se à sua concepção. Com relação à transferência de tecnologia para a indústria, um sistema desktop chamado "**P**layer **P**sychological **P**rofile Identification System", ou apenas 3PIS, foi desenvolvido contemplando o processamento de todas as características psicológicas propostas.

Palavras-chave: Comportamento de Jogador, Perfilamento de Jogador, Modelagem de Jogador, Ciclo de Vida de Uso de Jogos, Game Analytics, Perfil Psicológico de Jogador, Revisão Sistemática da Literatura, Player Psychological Profile Identification System (3PIS).

List of Figures

List of Tables

List of abbreviations and acronyms

ach	Achievement
aff	Affiliation
DB	Database
DBMS	Database Management System
DistH	Distance to Head
FFM	Five Factor Model
FN	False Negative
FP	False Positive
FPS	First-person Shooter
GRV	Game Refinement Value
GC	General Characteristic
GP	General Profile
GT	General Topic
HBM	Human-being's Model
HP	Health Points
ID	Identification Code
inf	Information
KDD	Knowledge Discovery in Databases
LSTM	Long Short-term Memory
mat	Materialism
MMORPG	Massively-multiplayer Online Role-playing Game
MOBA	Multiplayer Online Battle Arena
NPC	Non-player Character

NumP	Number of Players
PT	Personality Trait
PvE	Player versus
PvP	Player versus Player
pwr	Power
RI	Risk Indicator
RPG	Role-playing Game
RQ	Research Question Environment
RTS	Real-time Strategy
SVM	Support Vector Machine
sen	Sensual
TN	True Negative
TP	True Positive
UEF	Unification Explorer Framework
WoW	World of Warcraft
WOWAH	World of Warcraft Avatar History

List of symbols

$\veebar$	Exclusive OR
$\subseteq$	Subset
$\supset$	Supset
$\rightarrow$	Imply
$\nrightarrow$	Not imply
$\approx$	Approximate
$P(x)$	Probability of an event x happen

Contents

1 Introduction

Being able to understand the players' behavior is a valuable ability to game producers (also known as game organizers, game economy designers, product managers, or analytic game designers (GUARDASCIONE, 2018; TIMOTHY, 2020)) because together with a good comprehension of the active players comes the opportunity to offer new game contents that may better please them. Happy players tend to play longer, entailing in more profit to the gaming companies (KUMMER; NIEVOLA; PARAISO, 2017a; HARRISON et al., 2015).

A player starts to play for motivational reasons (a voluntary usage) and stops playing due to an absence of them (ZHU; LI; ZHAO, 2010; COOK, 2007; FAIRCLOUGH, 2008). The pleasure in play can be understood in many ways, such as the natural human behavior of a ludic seek (HUIZINGA, 2000) or the pleasure of having the needed ability to solve a problematic challenge (NAKAMURA; CSIKSZENTMIHALYI, 2009; CZISIKSZENTMIHALYI, 1990; HUNT, 1963).

Games may have different business models, such as subscription, pay-to-play, free-to-play, and sometimes, combinations of them (KUMMER; NIEVOLA; PARAISO, 2017a; SPELLER_III, 2012). Each business model demands different approaches by game producers. For example, pay-to-play games (shelf-games; from virtual or physical stores) usually do not have the capacity of providing new game content to their players after the game release, so the game producer has only one "shot" to try to entertain the players and obtain profit. On the other hand, subscription and free-to-play games are usually online games that allow the addition of new game content (the idea of Game as a Service (GaaS) (CLARK, 2014)), and consequently, the management of the game usage lifecycle (also known as game product lifecycle). This management aims at postponing the end of the lifecycle as much as possible, improving the income by doing so (MCALOON, 2018b). The usage lifecycle of a game starts with its first usage and ends when it is not more profitable, or there are no active players anymore (KUMMER; NIEVOLA; PARAISO, 2017a).

The game usage lifecycle management uses as a main base of information the called usage data (or telemetry data). This kind of data can have different granularities and portrays players' actions, events, or status in-game. With this kind of information in hand, game producers gauge the players' motivation by extracting and evaluating the called usage metrics[1] (hereafter, the terms "metrics" and "features" are used interchangeably).

[1] Even though all the metrics obtained from usage data can be referenced as usage metrics, this work assumes, hereafter, a specific nomenclature regarding the metric generation process. All the traditional metrics that are built without the support of any psychological basis are named as raw metrics, whereas the ones that use psychological information are named as psychological metrics.

Some examples extracted from (SPELLER_III, 2012) are shown next :

- MAU (Monthly Active Users):

$$\text{MAU} = \sum_{i=1}^{n} Pi \tag{1.1}$$

- WAU (Weekly Active Users):

$$\text{WAU} = \sum_{i=1}^{n} Pi \tag{1.2}$$

- DAU (Daily Active Users):

$$\text{DAU} = \sum_{i=1}^{n} Pi \tag{1.3}$$

- Sticky Factor per month:

$$\text{StickyFactorMonth} = \frac{DAU}{MAU} \tag{1.4}$$

- Sticky Factor per week:

$$\text{StickyFactorWeek} = \frac{DAU}{WAU} \tag{1.5}$$

- Profit;
- New players rate;
- Abandonment rate or Churn rate (as in some games there is no deregistration process, a policy to define when a player churned is needed, examples can be found at (LEE et al., 2018; KIM et al., 2017; TAMASSIA et al., 2016; PERIÁÑEZ et al., 2016; RUNGE et al., 2014));
- Retrieval rate (a player who abandoned the game and then returned to play again; an approach to retain returning players can be found at (MANSELL, 2015)).

Where $P_i = 1$ when a player $_i$ played a game during a given time-span and $P_i = 0$ otherwise; n is the total number of players. In Figure 1 it is possible to see the different behaviors of the MAU metric.

In the same way that it is crucial to understand the players' current behaviors, it is also essential to predict their future. This fact is highlighted by increasing investments of game companies, that follow the GaaS policy, to comprehend players even more deeply after the game release (KERR, 2018b; WAWRO, 2018; MCALOON, 2018c; KERR, 2018a). An academic research field that aims to identify, model, and predict the players' behaviors

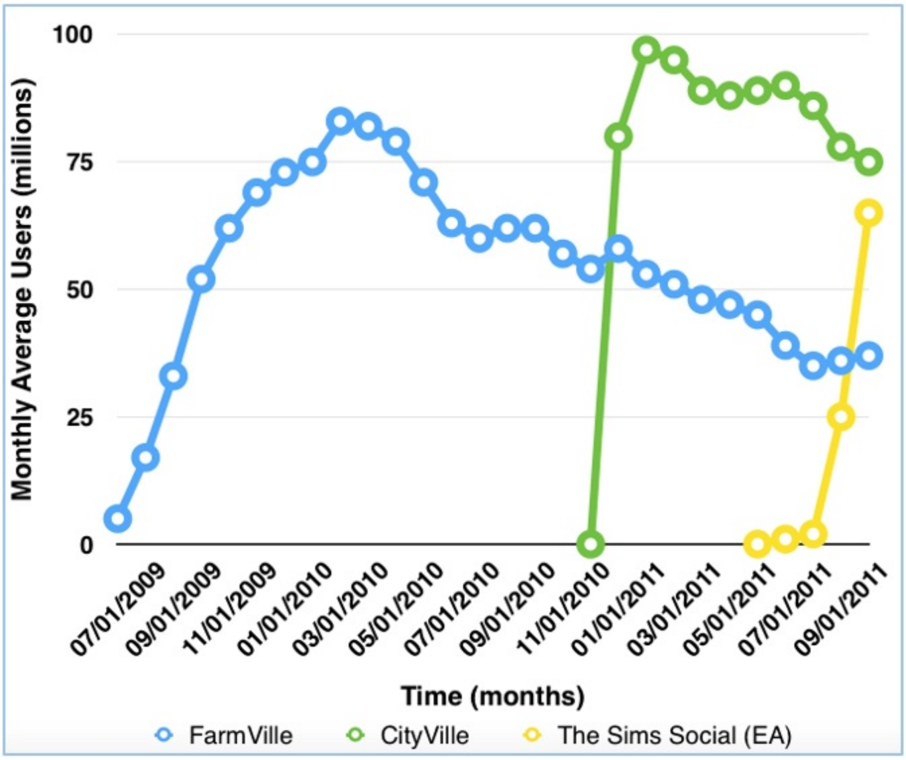


Figure 1 – MAU for some games, adapted from (SPELLER_III, 2012)

is the Game Analytics field (EL-NASR; DRACHEN; CANOSSA, 2016; WEBER, 2018; ZOELLER, 2010). This research area evolved in the same market direction[2] to deal with the new GaaS challenges, such as the management of churn, retention, monetization, bot detection, matchmaking, recommendation systems, new content testing, among others. As a field in evolution, the Game Analytics approaches can be divided into the following categories (DRACHEN, 2018):

- Snapshot: focuses on identifying players' profiles of a match (DRACHEN et al., 2012; DRACHEN et al., 2014).
- Dynamic: refers to the players' historical behavior and has a bigger scope than the snapshot. For example, the changes in usage time and in-game activities (NAYING et al., 2020; CASTRO; TSUZUKI, 2015; PERIÁÑEZ et al., 2016; BERTENS; GUITART; PERIÁÑEZ, 2017; LEE et al., 2018; GUITART; RÍO; PERIÁÑEZ, 2019; RÍO; GUITART; PERIÁÑEZ, 2020; KRISTENSEN; BURELLI, 2019; KUMMER; NIEVOLA; PARAISO, 2018a; JANG; KIM; YU, 2019; ROTHMEIER et al., 2020), and the identification of the current game lifecycle stage (KUMMER; NIEVOLA; PARAISO, 2018b) are considered. This approach is usually associated with the lifecycle stages and maps players' behavior changes over time.

[2] A survey containing the last 14 years of Game Analytics applications can be found at (FERNANDES; CASTANHO; JACOBI, 2018). Another exciting review that portrays the technical challenges to implement players' behavior prediction can be found at (HARRISON et al., 2015).

- Contextual: addresses the game's external data, for instance, geographic, demographic (e.g., gender, residence, educational level, income), and players' opinion data (HASSOUNEH; BRENGMAN, 2014; BARKER, 2009; PARK; KEE; VALENZUELA, 2009; LENHART et al., 2007; YEE, 2006).
- Spatio-temporal: this approach aims at clustering players' profiles based on the way that they explore 2D or 3D environments (DRACHEN; SCHUBERT, 2013; AUNG et al., 2019; MELO et al., 2019; MELO et al., 2020).
- Predictive: consists of using historical data to predict future behavior (SPELLER_III, 2012; BERTENS et al., 2018). For example, the identification of possible future income from a specific region (GUARDASCIONE, 2018).
- Lifetime: consists of looking at the remaining time (in-game) and motivations of players to generate metrics, for instance, the customer lifetime value (CLV) (MOREIRA et al., 2017), Average Revenue Per User (ARPU), Cost Per Install (CPI), and the monetization rate (i.e., the pace that players turn from free-users to paid-users) (CHEN et al., 2018; GUITART et al., 2019; GUITART; RÍO; PERIÁÑEZ, 2019; RÍO; GUITART; PERIÁÑEZ, 2020).
- Psychographic: this approach focuses on identifying the players' motivations and personalities.
- Psycho-predictive: the objective is to identify the current and future psychological state of players.

In part, Game Analytics works are restricted because it is not easy to obtain usage data as it is a confidential and critical information for game producers. Researchers usually obtain data throw agreements with game companies (SIFA et al., 2013; SPELLER_III, 2012; AUNG et al., 2018), manual or automatic gathering (LEE et al., 2011), or in Artificial Intelligence competitions (LEE et al., 2018). However although the academic side has difficulties and limitations to work, the game producers' side is putting more and more efforts to improve their internal analysis of players' behavior, highlighting the importance of this subject to business (KERR, 2018b; WAWRO, 2018; MCALOON, 2018c; KERR, 2018a).

Focusing on the Psychographic and Psycho-predictive categories, they can be divided into two main research lines: invasive and non-invasive. While invasive approaches demand the use of devices or questionnaires to infer psychological aspects of players (CALVO et al., 2015; WANG et al., 2019), non-invasive approaches are incipient and aim at automatically identifying psychological aspects of players using only the available telemetry data, facilitating the analysis of a higher number of players compared to the invasive approaches. This thesis is focused on the non-invasive perspective.

Non-invasive approaches are firmly attached to the knowledge provided by psychological models, following the ground concept of affective loop (SUNDSTRÖM, 2005; CALVO et al., 2015) and the balance between model-based and model-free approaches (YANNAKAKIS; TOGELIUS, 2011). Some initial works try to identify players' psychological profiles based on usage data (KUMMER et al., 2019; ODIERNA; SILVEIRA, 2019) as well as others successfully used the psychological descriptions to improve the identification of churn, survival time, and monetization candidates (JEON et al., 2017; BONOMETTI et al., 2019); however, the most common approaches target "home-made" games (LANKVELD, 2013; MAKANTASIS; LIAPIS; YANNAKAKIS, 2019), rather than commercial ones (PEDERSEN; TOGELIUS; YANNAKAKIS, 2009; SHAKER et al., 2011). It is interesting to highlight that, except by the psychological categories, Game Analytics approaches often focus on one game at a time and usually model players based on this specific game, rather than using more abstract models. It seems that each researcher has a different point of view about what constitutes players' behaviors, resulting in a diverse range of interpretations. In sum, it is possible to suggest that the success of non-invasive psychological approaches relies on the joining of two previous parallel research fields: the traditional Game Analytics and psychology. On the one hand, the traditional Game Analytics approaches propose metrics attached to the available usage data; on the other hand, the psychological perspective creates players profiles that are conceived without this data restriction. Therefore, this success arises from the proposition of enhanced metrics that contain psychological constructs. It is essential to highlight that manually linking data to psychological descriptions directly (i.e., without the support of questionnaires, inventories, or a systematic procedure) demands several cares and assessments that must be taken into account to avoid possible biases, such as the concepts of reliability and validity described by (YANNAKAKIS; COWIE; BUSSO, 2017).

Despite the success of the non-invasive approaches, they present three limitations. The first one regards the number of psychological models adopted, which is, as far as our knowledge goes, usually one. Linked to it, the criteria adopted by different authors to chose one or another psychological model lacks a systematic procedure to justify it. Given this, in a situation where a person wishes to take advantage of a model to try to solve a specific problem in the industry or academia, how can this person be sure that the chosen model is the most appropriate? Does this model depict all aspects of this person's needs? The second limitation regards the automatically and individually identification of players' psychological profiles considering a massive number of players, and the third, the assessment of such generated profiles.

A possible solution studied in this thesis to solve the first limitation regards the identification of a general model that covers all the other models' aspects quantitatively (i.e., a unified model; hereafter, "general model" and "unified model" are used interchangeably). Given a unified model, it is possible to explore ways to systematically and automatically identify its enhanced psychological descriptions occurrences on telemetry data for a massive number of players, generating players' psychological profiles as a result (solving the second limitation). To solve the third limitation, the generated psychological profiles can be used in risk prediction problems, where if the psychological data supported a better prediction, they could be assessed as accurate. These three limitations are overcome in this thesis (more details are given in the end of this Section).

As shown further in this work, the use of a general model is worth compared to using a single model due to its wider coverage of concepts. The notion of "psychological profile" adopted in this work regards the composition of psychological features. For instance, if a psychological model portrays aspects of four kinds of emotions (e.g., hope, fear, joy, and distress) and it was possible to systematically identify these emotions occurrences on data for each player on each time-span, there will be, at least, four psychological features. These features can vary in complexity and can be joined to generate other metrics. As simple approaches, it is possible to represent the emotions occurrences as boolean values (Table 1), or a counting format (Table 2). It is important to not misunderstand the term "psychological profile" for "players' psychological profiles". In this work, when the term "psychological profile" is used, it regards a set of psychological features. By contrast, when the term "players' psychological profiles" is used, it regards the values of each psychological feature assigned to the players.

Table 1 – Example of a psychological profile with boolean psychological features

Time-span	Player ID	Hope	Fear	Joy	Distress
2020-01-13	1A2B88	True	False	True	False
2020-01-13	0021S	False	True	False	True
2020-01-14	1A2B88	False	True	False	False
2020-01-14	0021S	True	False	False	False

Table 2 – Example of a psychological profile with counted psychological features

Time-span	Player ID	Hope	Fear	Joy	Distress
2020-01-01	1A2B	80	12	1	0
2020-01-01	0021DA	55	207	0	1
2020-01-02	1A2B	10	2	1	0
2020-01-02	0021DA	300	287	0	0

The aforementioned proposition of using a general model to extract a psychological profile is formalized in the following problem statement: "Is it possible to identify psychological profiles encompassing multiple psychological models on digital games based on usage data?". In the context of this thesis, a positive answer to this problem statement means that a set of psychological features were successfully extracted from telemetry data according to the psychological constructs of a unified model.

Bearing in mind the problem statement, it is possible to notice the importance that psychological models have in this work, as they are the main source, together with the telemetry data, for all the analyses and discussions. To clarify what is and why the term “psychological model” is adopted in this work, it is first needed to look at the terms used by different authors. It is interesting to note that some authors name their group of representations as theory, model, or even archetype, whereas others present their findings without using any of these terms. It was opted to not account for such terms as the desired pieces of information are the players' behaviors descriptions (such as preferences, motivations, and status) contained inside of such groups. Hence, adopting the term “psychological model” is just a formality to consider all the different representations over the same abstract concept, the idea of a “container of psychological descriptions”.

Even though the game context presents a varied number of psychological models, human behavior has also been modeled in other contexts, such as economics, philosophy, neuroscience, and artificial intelligence (YANNAKAKIS; COWIE; BUSSO, 2017). Thus, it is possible that some players' behaviors are not delineated by current players' models but by models from other contexts. This fact highlights the importance of encompassing models applicable to games that were not initially proposed to focus on them. Next, some examples of psychological models linked to games are presented.

As digital games are softwares, there is a fundamental model that describes them, the software lifecycle model (MOORE, 1995). In Figure 2 it is possible to identify the relation between the number of active users of a software and time. In special, this model illustrates the called “The Chasm” that means the initial software acceptance. When a software is about to be released, there is an expectancy in the users' minds. If that expectancy is not fulfilled after first usage, then the user may abandon the software earlier than usual, ending the usage lifecycle abruptly. In a normal course of life, a software starts and finishes its usage lifecycle with its first usage and with its last usage or when the software is no more profitable. The usage of a software may reduce for several reasons, for instance, the release of a competitor's software, or the problem that the software solves does not exist anymore.

In the entertainment digital game perspective, the users (players) are guided by motivational reasons (a voluntary usage), moreover, there are different game genres, such as: RPG, MMORPG, MOBA, and RTS. According to (COOK, 2007), a game genre is defined by its game mechanisms, therefore games with similar mechanisms tend to be in the same genre. Cook defined the genre lifecycle as portrayed in Figure 3 where:

- Introduction (Intro): the game mechanisms are innovative and attract attention.
- Growth: the users have accepted the game (the crossing through "The Chasm" (MOORE, 1995)) and more games of this genre are produced.

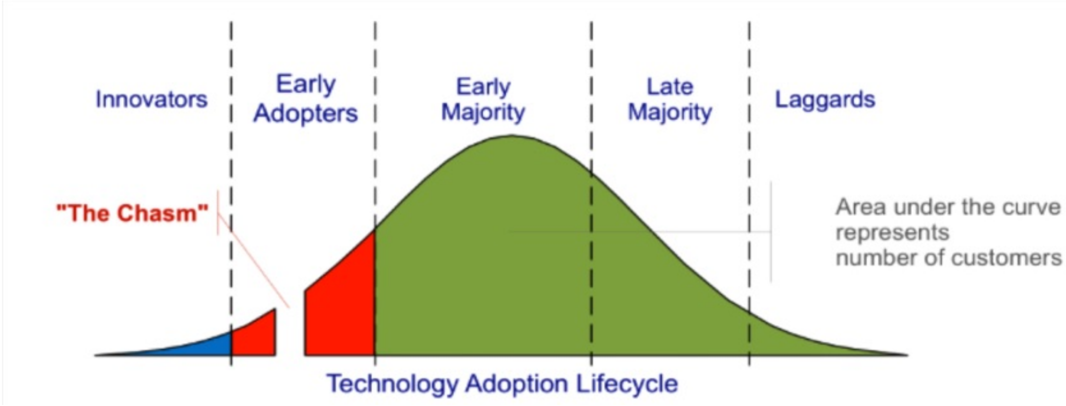


Figure 2 – Software lifecycle, extracted from (MOORE, 1995)

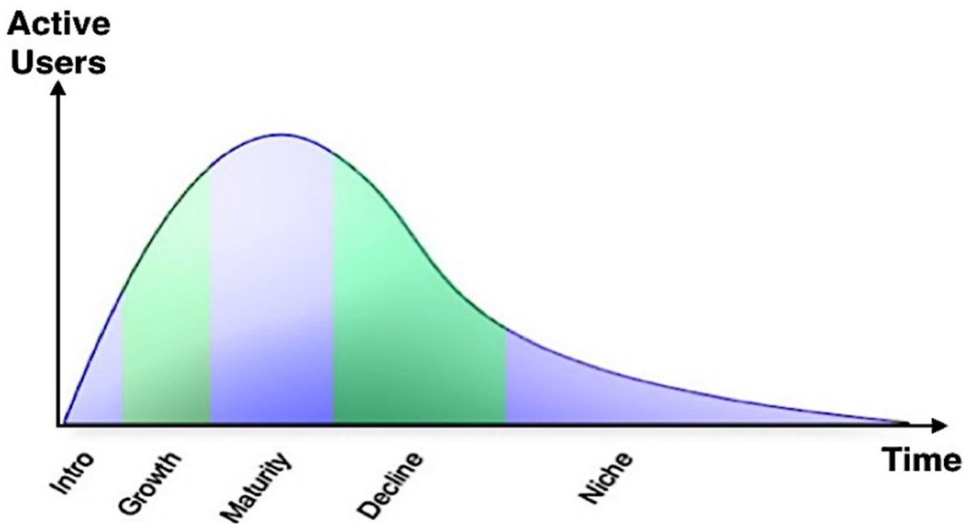


Figure 3 – Game genre lifecycle, adapted from (COOK, 2007)

- Maturity: great game producers adopt the genre. According to (CASVEAN, 2015), developing games in mature genres reduces the risks of an abrupt end of the game usage lifecycle, the idea of falling in “The Chasm” of Moore.
- Decline: fewer games are produced, the genre does not attract new players as they used to.
- Niche: there is no financial return, great game producers leave the genre, and some games are maintained for love and not for money.

According to (GARDA, 2013), new game genres can surge from game upgrades where the game mechanisms changed or were mixed with other mechanisms. A problem is that the new genre may not please the players of the previous version. New game genres may also surge with no base in other genres or games.

Besides the acceptance dynamics of the game genre’s lifecycle, an individual game also carries its specific notions of acceptance. Players can have different motivations to play the same game; for example, some prioritize friends and building, others prefer to explore the game environment, and others may seek combat challenges (VANDENBERGHE, 2018; HUNICKE; LEBLANC; ZUBEK, 2004; BARTLE, 1996; THUE et al., 2007). These differences mean that the players' motivations to play a specific game are firmly attached to the game content offered to them, suggesting that the more diversified a game content is, the more diversified its players' motivations will be. However, players' motivations can extinguish for several reasons, such as completion of all game challenges, the absence of new and exciting challenges, the end of social interactions in-game, lack of support from the game company, or achieving full mastery over all game mechanisms (ZHU; LI; ZHAO, 2010; COOK, 2007; FAIRCLOUGH, 2008). Given this, Cook also identified players' behaviors linked to the game lifecycle, such as:

- Players' stages

  Initial learning: an initial experience to evaluate the game mechanisms and challenges.

  Master: mastery over all game mechanisms.

  Tool: a game as a tool to achieve objectives.

  Burnout: loss of opportunity to play.

- Players' abilities

  New player: a new and fun experience (first interactions).

  Mature player: expert in the game mechanisms.

  Niche player: this kind of player can be understood from three different points of view:

  Fire keeper: the player does not give up the game and keeps playing.

  Lapsed player: a new life routine prevents the player from continuing playing as before, the player's abilities drop.

  Players with no network support: players who found the game for the first time with no references, but the game could be obsolete for a long time ago.

In the MMORPG (Massively-multiplayer Online Role-playing Game) genre, (ZHU; LI; ZHAO, 2010) identified four motivational stages of players:

- Try: when the player tries the game for the first time.
- Tasting: after the player's approval about the game, he/she starts to accumulate profit, such as friends, items, levels, and quests.

- Retention: in this stage, the player already knows all the game mechanisms and game content and starts to have a lack of interest in continuing playing, but stills to play because of friends.
- Abandonment: there is nothing that can change the player's mind about leaving the game.

It is interesting to notice that each psychological description of each model presents its own complexity, and because of that, the complexity of linking these philosophical descriptions to mathematically based forms (i.e., a psychological profile) also varies. For instance, identifying the "Retention" occurrence must encompass not trivial notions of social interactions degree of a player attached to his/her accomplishment over the game content challenges. By contrast, the "Try" occurrences are more straight forward, as it is just needed to check if the player is a newcomer (KUMMER et al., 2019).

To identify psychological models applicable to games, a Systematic Literature Review (SLR) is proposed in Chapter 2. This review identified 67 unique psychological models that were divided into two groups. One containing 46 models exclusive to the game context, and another with 21 regarding a general context of human behavior (i.e., not limited to a specific context). To identify a general model for each group, the Unification Explorer Framework (UEF) is proposed. This framework assesses the coverage of models by applying the concept of holism.

Holism was first coined by (SMUTS, 1927) as "the fundamental factor operative towards the creation of wholes in the universe" which caries the idea of "a whole is greater than the sum of its parts". Summing it up, holism provides a means to describe a context's capabilities by considering both the parts' capabilities and the parts interactions' capabilities. A capability can be seen as a result, a piece of information, or a possible operation. The holism concept is currently adopted by different contexts, such as social, biological, physical, or economical, to analyze the resulting value of systems (OSHRY, 2007; AUYANG, 1998). Applying it to our case, players' psychological models, a whole is a set of psychological models, the parts are the distinct models of this set, and the interactions are the sharings of characteristics between these models. Therefore, in this thesis, the idea of a unified model depicts a single view that contains all points of view from a set of psychological models (the context), considering also the knowledge present on the models' interactions.

The analysis of the resultant unified models from the game and general contexts showed that the general perspective presents a better linkage to the telemetry data and greater coverage of psychological concepts than the exclusive game one (solving the previously presented first state-of-the-art limitation). This fact guided the next step, the proposition of a method that generates a psychological profile of players based on usage data (solving the second state-of-the-art limitation). It means that the resultant profile of this work is based on a general view of the human behavior and not limited to the game context. An attractive quality of the proposed profile is that it encompasses psychological essences regarding short, mid, and long-term aspects, regarding the notions of human need, emotion, sentiment, and personality.

After generating the profile it is needed to validate it. However, the problem is that there is no true label to compare with as the only available information regards usage data. Given this, an alternative was suggested. To assess the proposed profile coherence with actual players' psychological essences, it was proposed to use the generated psychological metrics in a churn prediction problem. By comparing the features performance with the state-of-the-art approach (that adopts only raw features), it is possible to assess if the generated metrics makes sense, at least, in the churn context.

The proposed method was implemented in a desktop system named "**P**layer **P**sychological **P**rofile Identification System", or just 3PIS. The 3PIS processed a dataset of an MMORPG containing approximately 10,000 players (LEE et al., 2018). The generated profile was used in the proposed assessment procedure, which resulted in the approval of the proposed profile structure, as it obtained a better churn prediction result in comparison to the state-of-the-art (solving the third state-of-the-art limitation). This positive assessment also allowed the coining of the term "Psychological Data Enhancement", which represents the proposed method essential conception.

Besides interpreting players' behaviors through the lens of psychological models, another study was approached in this thesis regarding the idea of interpreting these behaviors by physics concepts applied to the human brain. This kind of modeling was initially proposed by Iida et al. (IIDA; TAKESHITA; YOSHIMURA, 2003) and named "The Game Refinement Theory". This thesis approaches some experiments where one of this theory resultant metrics, the Game Refinement Value (GRV), is joined with a raw metric (i.e., the Commitment metric[3]), point to exciting findings. Also, a joining between the proposed method and the GRV is proposed, representing a bridge between these two, until, parallel research approaches.

Moving back to the problem statement, it is possible to divide its answering process into three main parts: the first where the psychological models are approached, the second where the psychological profile of players is built by linking psychological construct to usage data, and the third where the generated players' profiles are assessed in a churn prediction problem. In addition to the problem statement, 11 research questions are proposed and answered, describing some key concepts about the retrieved works of the SLR and the literature meaning behind each psychological feature present on the proposed psychological profile.

[3] This metric is detailed further in this thesis.

Despite the improvements identified in the churn prediction, each psychological feature contained in the proposed profile has useful information, as each one represents a linkage between the textual descriptions of psychological models and a mathematical representation. By providing measures to the psychological descriptions, game producers can better understand players individually, as the proposed features highlight motivational aspects of them. Moreover, novel metrics were proposed to explore additional benefits of the psychological features linked to game companies' concerns. In particular, the linking between the notion of Game Path[4] and the churn label allowed a change of perspective in the churn management from reactive to proactive.

Even though the decision-making process done by game producers and game designers based on the proposed psychological features are not part of our scope, it is essential to highlight that the features proposed in this work can support them by giving answers to the following questions that are part of their daily life:

- What are the individual players' motivations for a given game?
- What is the degree of similarity between the active players' behavior in a game?
- Is a game entertaining its players with a comfortable/desirable degree of challenge?
- Was a given game upgrade successful? What was a possible cause?
- When should I release a game upgrade?
- What are the game design components that players most chase in a given game?
- What components of a game design should be added to increase engagement?
- What components of a game design should be removed or modified to increase engagement?
- What is the earliest moment when I can identify churn candidates?
- What would be an appropriate way to understand individual churn behaviors?
- What makes players start or continue playing?

All these questions are answered in the concluding Chapter of this thesis. In special, we highlight that all of these answers are obtained in a non-invasive way, not needing to interact with the players.

Besides the support for game designers and game producers, we wish that this work inspires researchers in finding new ways of extracting useful knowledge from psychological

[4] As explained further, a Game Path regards a structure that contains shared players' choices in-game, represented by their sequence of actions.

models or human brain nuances, and also in applying this novel knowledge to the varied challenges present in the broad game context and beyond.

1.1 Objective

The proposition of a method capable of identifying the current psychological profile of players through a machine learning approach over usage data.

1.2 Problem Statement and Research Questions

To achieve our objective the following problem statement is proposed.

- **Problem Statement:** *Is it possible to identify psychological profiles encompassing multiple psychological models on digital games based on usage data?*

To answer the problem statement we propose 11 research questions (RQs).

- **RQ1:** *What is a psychological aspect?* As a psychological aspect is a vague term due to the human-being complexity. We wish to describe it in more specific terms.
- **RQ2:** *What is a psychological profile?* This questions aims at describing the concept of psychological profile and establishing its linkage to psychological models.
- **RQ3:** *What are the psychological models applied to games?* We wish to list the psychological models that were somehow applied to games.
- **RQ4:** *Is it possible to link a profile of one model to the profile of another model?* The idea of this question is to check if different models deal with the same aspects.
- **RQ5:** *Can psychological models be ranked?* To select a model, a way to judge which model is more appropriate is needed. Here, we wish to investigate how a model can be compared to another.
- **RQ6:** *Is it possible to combine models?* The idea is to validate if more accurate models can be created by combining different ones.
- **RQ7:** *Is there a general psychological model that can portray all the players' aspects?* This research question aims at assessing if it is possible to use only one model to represent all the players' aspects.
- **RQ8:** *Are all models applicable to all game genres?* We want to check if the psychological models can be applied independently of game genres.

- **RQ9:** *What are the advantages and disadvantages of using psychological models?* We want to identify these advantages and disadvantages and list them, associating real applications if possible.
- **RQ10:** *To what extent characteristics of usage data can be used to identify psychological profiles?* As human-beings' psychological aspects are complex, we want to identify to what extent these characteristics can be identified based on usage data.
- **RQ11:** *How an identified profile on usage data can be assessed?* We wish to identify ways to validate the identified profiles.

1.3 Research Methodology

In this research, we use two methodologies to answer the research questions, namely (1) Systematic Literature Review and (2) Development Research (MAREN, 1996). The first was applied to identify a list of psychological models applied to games. The second was used to propose the Unification Explorer Framework and the method that identifies psychological profiles. Additionally, experts' assessments (from psychologists) were performed. In sum, there are four research topics were the research questions are approached, such as depicted by Table 3. As we are not psychologists, all the psychological concepts were validated with experts [5].

Table 3 – Research questions per research topic

Research Questions	Systematic Literature Review	Unification Explorer Framework	Method Proposition	Expert's Assessment
RQ1	X			X
RQ2	X			X
RQ3	X			
RQ4		X		X
RQ5		X		X
RQ6		X		X
RQ7		X		
RQ8	X	X		
RQ9	X	X	X	
RQ10			X	
RQ11			X	X

[5] Prof. Dr. Tatiany Honorio Porto Aoki and Prof. MSc. Ulisses Domingos Natal, thank you for your careful assessment. Giving more details about this assessment process, the author of this thesis presented to the experts his interpretations about the psychological aspects and the proposed connections between them. The experts checked each interpretation by pointing the flawed understandings and providing the correct interpretation, and validating the proposed philosophical connections between these aspects. Besides, additional literature was suggested to support a better understanding of the fundamental pillars of psychology.

1.4 Contributions and Technology Transfer

We can summarize the work contributions into nine aspects, namely:

- Scientific Contributions

  A list of psychological models applied to games.

  A framework that describe means to compare and rank psychological models (i.e., the UEF).

  A summary of advantages of using psychological models.

  The proposition of a method capable of identifying the current psychological profile of players.

  The accuracy improvement of models that predict risk situations in the Game Analytics field.

  A change of perspective in the churn management from reactive to proactive.

  A set of psychological features that measures philosophical descriptions, supporting game producers with a more prosperous point of view about what constitutes their players' motivations.

- Technical Contributions

  The 3PIS.

  The possibility to apply the proposed method in other contexts.

The proposed method can be applied in real situations and in other contexts besides the game one, where people's likes and dislikes could be identified based on data.

1.5 Scope

The proposed method to identify players' psychological profiles is initially applied to a massively multiplayer online role-playing game (MMORPG). The method starts analyzing the usage data and ends labeling players with psychological profiles. The data collection procedure and decisions made based on the provided psychological information are not part of our scope, even though they are described to better clarify the application environment.

1.6 Organization of the Text

This work is organized as follows: in Chapter 2 the SLR is presented, and its research questions are answered; in Chapter 3 additional aspects regarding MMORPGs, usage data, Game Path, risk situations, Game Refinement Theory, Knowledge Discovery in Databases (KDD), the Commitment metric, and Concept Lattices are presented; the Unification Explorer Framework is presented and applied considering the two sets of psychological models identified in the SLR in Chapter 4. Also, its linked research questions are provided; in Chapter 5 the method to extract psychological features from usage data is proposed, and some research questions are answered; Chapter 6 presents the method application to an MMORPG together with analysis, discussions, and comparison (with the Game Refinement Theory); next, in Chapter 7 the resultant psychological profile is assessed in a churn prediction problem; and finally, conclusions and future works are pointed in Chapter 8. Given this, the first main part of this thesis (i.e., the psychological models identification and analysis) are approached until Chapter 4, the second main part (i.e., the proposition and analysis of a psychological profile of players) regards the Chapters 5 and 6, and the third main part (i.e., the profile assessment) is presented in Chapter 7.

2 Systematic Literature Review

The main concern about this review is to provide a guide of psychological aspects in games, but psychological aspects is a vague term, and we had to put some extra effort into understanding better and illustrate it.

Our starting point was based on the work of (TINBERGEN, 1951)(Nobel Prize winner of Physiology or Medicine in 1973), which elucidated the behavioral aspects of animals. In his work, these aspects were divided into four categories of explanation (based on Aristotle's four cases (CHARLTON et al., 1983)) such as follows:

- Population aspects

  Function (adaption): a species trait that solves a reproductive or survival problem in the current environment. This category is based on Aristotle's final cause, which regards "the objective of something, the sake for which a thing is done".

  Phylogeny (evolution): the history of the evolution of sequential changes in a species over many generations. It is linked to Aristotle's formal cause "the pattern or form that allows the identification of something".

- Individual aspects

  Mechanism (causation): mechanistic explanations for how an organism's structures work. It is related to Aristotle's efficient cause "an interaction to produce something; changing an object shape for the desired purpose".

  Ontogeny (development): developmental explanations for changes in individuals, from DNA to their current form. This last category is linked to Aristotle's material cause "the nature of the raw material out of which an object is composed".

Figure 4 illustrates the relations between those categories. On the left side are the population aspects, and on the right side, the individual ones. In special, we are interested in the behavior subject that is caused by the mechanism category, as this category embraces the psychology (TINBERGEN, 1963; TINBERGEN, 1951).

Inside psychology, there are different approaches; in special, we give attention to two of them, the psychoanalytic and cognitive ones. On the one hand, the psychoanalytic idea points that humans are so complex that each person is unique (a subjective approach; philosophical)(FREUD, 1912; FREUD, 1920), so it is not possible to model them. On the other hand, the cognitive approach accepts that each person is unique, but it does not prevent the fact that different people may behave similarly, so in this case, it is possible

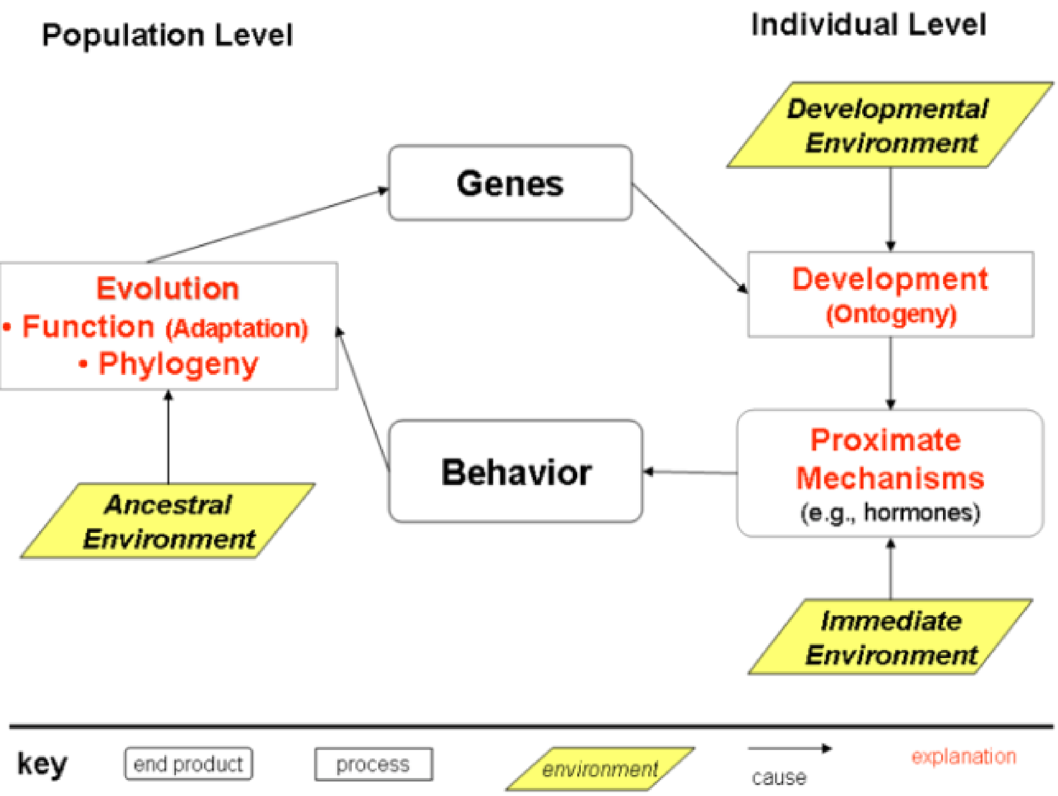


Figure 4 – Causal Relationship, adapted from (TINBERGEN, 1963)

to model people based on what they do (a scientific approach; behavioral)(SKINNER, 2003; MOREIRA; MEDEIROS, 2018). Therefore, this research is situated on the cognitive aspect of psychology that is comprehended by animal behavior's mechanism category.

There are many terms used in psychology, such as emotions, sentiments, opinions, personality traits, human needs, instincts, competence, and affect. The works of (MUNEZERO et al., 2014; CARVER; SCHEIER, 2012; KLEINGINNA; KLEINGINNA, 1981) aim at solving possibles misunderstandings about these terms. In this work, we deal with affect, emotions, sentiments, personality, personality traits, competencies, and human needs according to the following representation:

- Affect: it exists outside of consciousness before the development of awareness (MASSUMI, 1987); an affective matrix linked to a predisposition of the bodily systems to react in a certain way to internal or external stimuli, a higher-order category under which both feelings and emotions fall (MATTHIS, 2000).
- Emotion: an internal human feeling regarding some fact, with or without interactions with other humans (it is considered to have a short-term influence over people' actions)(MUNEZERO et al., 2014). For example, the Ekman's six basic emotions (EKMAN, 1992): anger, disgust, fear, happiness, sadness, and surprise.
- Sentiment: it is a positive, negative or neutral polarity regarding some fact or wish

(it is considered to have a mid-term influence over people's actions) (BEN-ZE'EV, 2000; RUSSELL; BARRETT, 1999; MUNEZERO et al., 2014).

- Personality: it is defined as the stable pattern of variation in individual acting, thinking, and expressing (LANKVELD, 2013).
- Personality Trait: it represents how a person behaves in different situations (it is considered to have a long-term influence over people's actions) (CARVER; SCHEIER, 2012).
- Competence: similar to personality traits, competence points a person's skills to overcome obstacles, improve him/herself, and identify his/her own being (GOLEMAN, 1998).
- Human need: it can be divided into two great groups, the physiological needs, and the psychological needs (MURRAY, 1938; MASLOW, 1968). Attaining a need is an achievement (physical or mental). A physiological need can be hunger, thirst, sleepiness, among others. While a psychological needs can be a social relation, sense of power, sense of belonging, sense of possession, sense of control, among others (CARVER; SCHEIER, 2012). To fulfill (attain) a physiological need, a human can drink water, and to fulfill a psychological need, make friends for example.

Regarding the identification of psychological models, a previous SLR (BOYLE et al., 2012) that identified them was extended with the proposed SLR of this thesis. Next, details of the proposed SLR regarding the search databases, the adopted keywords, the period of search, and the criteria of acceptance and rejection are presented. The summary of the retrieved works is also shown. After, the related works are presented and then the SLR related RQs are answered.

2.1 Search Databases

The following databases were used in this research: ACM Digital Library, IEEE Xplore, Science Direct, Springer Link, AAAI, Google Scholar, and Gamasutra. The databases of ACM Digital Library, IEEE Xplore, Science Direct, Springer Link, and AAAI were chosen because they focus on computer science, but also have conferences focused on games and journals focused on the psychological aspects between humans and computers. Google Scholar was chosen due its contents that include papers, books, master's and doctoral theses, and patents. Moreover, it also serves as a link between many databases from different research fields, so if we found a new database in the Google Scholar results, we could add it to the list of search databases for this thesis (the presented list is the final one). The Gamasutra is not an academic database, but we kept it as a point of access to the industrial approaches, as it has news about games and game producers, highlighting commercial aspects, players' feelings, game producers' concerns, and sharing of new problems and solutions. The databases of IEEE Xplore, Springer Link, AAAI, Google Scholar, and Gamasutra were not approached by the previous review.

2.2 Keywords

The previous review considered the set of the following keywords:

- ("computer games" OR "video games" OR "serious games" OR "simulation games" OR "games-based learning" OR "MMOG" OR "MMORPG" OR "MUD" OR "online games")
- ("evaluation" OR "impacts" OR "outcomes" OR "effects" OR "learning" OR "education" OR "skills" OR "behaviour" OR "attitude" OR "engagement" OR "motivation" OR "affect")

We first explored the wider game context in which the psychological models exist. We started with the game usage lifecycle concept, as it considers general aspects of interactions and motivations between a game and its players (our previous SLR (KUMMER; NIEVOLA; PARAISO, 2017a) was used as a main source of knowledge). Next, we searched for literature regarding player modeling. After this reading, we identified some conferences and journals that focus on the psychological aspects of players (e.g., motivations, preferences, and status). We read the papers from those sources and found our main keyword: "Player Psychological Profile". This keyword presented useful, allowing the identification of models not present in the previous review, as shown in Table 4.

Given the aforementioned literature exploration, it was observed that the concept behind the keyword "Player Psychological Profile" encompasses, in most cases, an abstract perspective by considering the terms profiling and modeling as synonymous to represent the process of textually describing common players' behaviors, which is a desired aspect to this thesis. However, such terms can have different meanings in specific research areas. In some cases, profiling is assumed as a top-down perspective that identifies shared behaviors between players (i.e., motivations, desires, preferences, status, engagement, and enjoyment). In contrast, modeling is seen as a bottom-up perspective, where individual aspects are preferred instead of the common ones. Also, modeling has another meaning in Machine Learning approaches (EL-NASR; DRACHEN; CANOSSA, 2016) by referencing the prediction models, which can be built based on players' profiles. The simulation of players is another research area with its idea of modeling, which regards the translation of players' profiles to actions in-game (COWLEY, 2009). Bearing in mind this thesis's aim in using the holism concept to emphasize the generality aspect of psychological models, the adoption of the keyword "Player Psychological Profile" is the correct way to go. It is justified by the fact that this keyword encompasses on both abstract and specific perspectives, the literature focus on describing shared players' behaviors, which potentializes the holism findings. In particular, the findings regarding the parts interactions' capabilities, which grows according to the sharings of common behaviors between different models (more details are presented in Chapter 4).

It is important to not misunderstand the varied meanings of player modeling with players' psychological models. In this work, players' psychological models are "containers" of textual descriptions of players' behaviors, without any regard to the manner utilized to build them (Section 2.6.1 presents more details about the structure of psychological models).

2.3 Period of Search

Even though the collection of usage data probably started in 1999 (WEBER, 2018), with the first big Game Analytics application in 2009 (ZOELLER, 2010), we opted not to limit the search period as players profiling is not limited to digital games or usage data. Another factor that influenced our decision was the fact that psychological models have references from the 19th and 20th centuries that are still accepted as state-of-the-art (ALLEN, 2015).

2.4 Criteria of Acceptance and Rejection

We included any literature that was linked to the players' psychological profiles regarding entertainment games. For example, if textual descriptions of players' behaviors were portrayed as results of analysis or predictions, or an application regarding the psychological aspects of players was presented, the literature met our criteria. We did not include literature if it was about the game development lifecycle, game design, serious games, or did not have any information regarding the psychological aspects of the game context.

2.5 Summary of Found Works

After applying the proposed protocol, a total of 315 papers were found and analyzed. Excluding the cases were no links to psychological aspects were identified in the abstract and in the usage of the adopted keyword, full readings were performed, totalizing 109 accepted papers according to the criteria of acceptance and rejection. Table 4 describes the number of papers found and accepted as well as the number of identified psychological models regarding all considered reviews. It is possible to notice that the proposed SLR extended the coverage of the previous SLR as new psychological models were identified, summing 46 unique models of players' behaviors (COOK, 2007; ZHU; LI; ZHAO, 2010; BARTLE, 1996; BATEMAN; BOON, 2006; THUE et al., 2007; BULATOV, 2018; FAIRCLOUGH, 2008; BRAYSHAW; GORDON, 2016; SMITH, 1997; PINE; GILMORE, 1999; KELLAR; WATTERS; DUFFY, 2005; YEE, 2006; MALONE; LEPPER, 1987; WEILLER, 2015; KOSTER, 2005; LAZZARO, 2004; COLWELL, 2007; VANDENBERGHE, 2018; HUNICKE; LEBLANC; ZUBEK, 2004; CAMMARATA; KOSTER; Google's Advanced Technology and Projects (ATAP) group, 2018; NACKE; BATEMAN; MANDRYK, 2014; CAILLOIS, 2001; YEE; BAILENSON, 2007; HARBORD; DEMPSTER, 2019; JERČIĆ, 2019; REILLY; ROOY; ANGUS, 2019; DEMETROVICS et al., 2011; O'BRIEN; TOMS, 2008; MARCZEWSKI, 2015; SHERRY et al., 2006; CHOU; TSAI, 2007; KIM; ROSS, 2006; FROSTLING-HENNINGSSON, 2009; LIN; LIN, 2011; JANSZ; AVIS; VOSMEER, 2010; OLSON, 2010; EGLESZ et al., 2005; GRIFFITHS; DAVIES; CHAPPELL, 2004; KOO, 2009; PRZYBYLSKI et al., 2009; WU; WANG; TSAI, 2010; HSU; LU, 2004; BLACOW, 1980; LAWS, 2002; DICKEY, 2007; HAGGIS-BURRIDGE, 2020).

A snowballing was applied to the obtained literature, which resulted in the identification of 21 general psychological models that are not focused on players (i.e., not limited to the game context), where some of these models were improved over time (MYERS et al., 1998; TOPRAC; ABDEL-MEGUID, 2011; FRIJDA, 1986; SMITH; ELLSWORTH, 1985; EKMAN, 1992; MEHRABIAN, 1980; MEHRABIAN, 1995; MEHRABIAN, 1996; RUSSELL; MEHRABIAN, 1977; HUIZINGA, 2014; DECI; RYAN, 1985; DECI; RYAN, 1995; MASLOW, 1968; MURRAY, 1938; ZILLMANN, 1988; ZILLMANN, 2015; ZILLMANN, 1995; ZILLMANN, 1996; ORTONY; CLORE; COLLINS, 1990; PLUTCHIK, 1980; NAKAMURA; CSIKSZENTMIHALYI, 2009; CZISIKSZENTMIHALYI, 1990; HUNT, 1963; GOLEMAN, 1995; GOLEMAN, 1998; GOLDBERG, 1990; EYSENCK, 1967; EYSENCK, 1973; NORTHROP, 1974; NORTHROP, 1984). We named these models human-being's models (or HBMs), and similarly to what happened to the players' models, the proposed review identified models not covered by the previous one, as also shown by Table 4.

Given the two identified contexts of psychological models applied to games, where one is specific to the players' behaviors, and the other presents a general perspective of the human-being behavior, regardless of context, the analysis from now on will be segregated for each perspective. It was opted because we consider players' models more specific than human-being's models. In addition to it, we assume that all players' models have an inherit link to all human-being's models, because, first of all, players are humans. Nevertheless, later in Chapter 5, a comparison between the two perspectives concerning its linkage to usage data is presented.

Table 4 – Summary of found works

Subject	Papers found	Papers accepted	Players models found	HBMs found
Previous SLR	19,776	55	16	11
This SLR	315	109	32	12
Snowball	103	39	10	21
Total	20,194	204	46(unique)	21(unique)

It is essential to highlight that the total number of psychological models is unknown as far as our knowledge goes. Therefore, in this work, we focused on psychological models that were somehow linked to games and players.

2.6 Retrieved Works

This section focuses on presenting all of the retrieved models through the lens of the proposed hierarchical structure of psychological models. Firstly, this hierarchical structure is explained, and then the next two subsections present the retrieved players' models and human-beings models, respectively.

2.6.1 Psychological Model Hierarchical Structure

Psychological models are textual descriptions of human behaviors, where different approaches present different degrees of abstraction. Thus, to allow for a comparison, the models must be seen in the same format. After analyzing the 67 identified models, a structure was proposed (Figure 5).

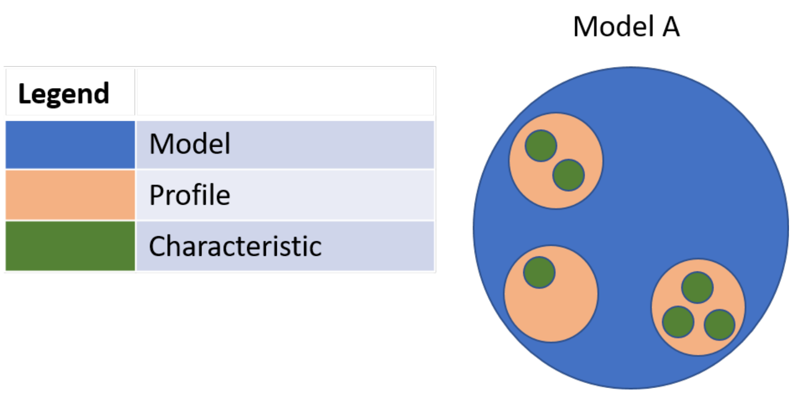


Figure 5 – Proposed psychological model hierarchical structure

In this structure, "Model" refers to the model's name (which is assumed as the model's authors' names together with its reference), "Profile" is the name given by the model's authors to a specific behavioral pattern, and "Characteristic" is a detailed explanation of a "Profile." Note that there is no limit to the number of profiles in a model or the number of characteristics in a profile. In summary, the idea of hierarchy is represented by the fact that the most prominent component, the model, has smaller parts, the profiles, which in turn also have smaller parts, the characteristics. It is also possible to represent this structure in a table format, as depicted by the example in Table 5.

Table 5 – Example of the hierarchical structure of a model in a table format

Model	**Profile**	**Characteristics**
Authors + Ref	Painter	Players who like to change the environment colors
	Fisher	Players who prefer to go fishing in saltwater or freshwater, with other persons or alone

This hierarchical structure was adopted because it was the most similar format across all of the models, which means that some models already followed this format while others needed some adjustments. We understand that this process of formatting models to the proposed structure is susceptible to personal bias and demands additional cares to mitigate it. Using the earlier example, one could point out that the Painter profile has only one characteristic, the fact that "a player likes to change the environment colors", and that the Fisher profile has two characteristics, the fact that "a player prefers to go fishing in saltwater or freshwater" and "a player prefers to be with other persons or alone". However, another person may divide the Fisher profile into four characteristics: "prefers fishing in saltwater", "prefers fishing in freshwater", "prefers to be alone", and "prefers to be with others". Moreover, another could consider a conditionality between the desire to be or not with others and the fishing activity. As we can see, different divisions may occur depending on personal interpretations, and as a manner to mitigate this bias, it is suggested splitting profiles descriptions into characteristics by looking at key points, such as an action, a quality, or a status, which portray what a person does, is, has, or desires, also contemplating the conditional aspects. By applying this in the previous example, "fishing" can be seen as an action and "being or not with others" as a desire, and assuming a condition between them, a single characteristic would be proposed. In conclusion, even though there is a certain degree of bias during the interpretation of models, the critical factor to retaining reliability is the maintenance of the original characteristics' meaning, regardless of the number of divisions. The analysis of this thesis uses this guideline to mitigate the bias problem.

2.6.2 Players' Models

All retrieved players' models were read, summarized, and formatted in the proposed hierarchical structure (table format). All descriptions can be found in Table 6. The term "environment" refers to the game environment and its possible interactions. The characteristics column presents players' status, preferences, motivations, and the general idea of playing (i.e., general descriptions of the game context). Also, this column was not segregated into distinct characteristics to allow the reliability assessment proposed by (YANNAKAKIS; COWIE; BUSSO, 2017) per future researches. Regarding cases where different models shared the same profile name, the models' authors' names were added to the profile names to differentiate them. The symbol "**" indicates that the profile name was not given by the original authors. To improve the layout of Table 6, the models references were not inserted. The map for each Model name and its reference is presented in Table 7.

Table 6 – Identified players' models

Model	**Profile**	**Characteristics**
Cook	Introduction	The generation of interest
	Growth	Growth of acceptance
	Maturity	Game well accepted
	Decline	Initial loss of interest
	Niche	End of interest and low acceptance
	Initial Learning	An initial experience to evaluate the game mechanisms and challenges
	Master (Cook)	Mastery over all game mechanisms
	Tool	A game as a tool to achieve objectives
	Burnout	Loss of opportunity to play
	New Player	A new and fun experience (first interactions)
	Mature Player	Expert in the game mechanisms
	Niche Player	Fire keeper; lapsed player; and no network support
Zhu et al.	Try	First try in a new game
	Tasting	Player's approval and accumulation of profits (e.g., friends, items, quests, etc.)
	Retention	Mastery over game mechanisms, completion of all challenges, loss of interest, and positive social interaction
	Abandonment	Loss of interest and absence of friends
Bartle	Achiever (Bartle)	Acquisition of rewards
	Explorer (Bartle)	Exploration of the virtual world

To be continued

Model	**Profile**	**Characteristics**
	Socializer (Bartle)	Positive social interactions
	Killer (Bartle)	Negative social interactions, preference for combat and interference in others' gameplay
Bateman and Boon	Conqueror	Achievements of all game challenges and recognition
	Manager	Solving of problems, proposition of strategies, and seeking to develop skills
	Wanderer	Fun experience attached to escapism (i.e., leave behind concerns of the daily life)
	Participant	Positive social interactions as a member of a group
Thue et al.	Fighters	Preference for combat and aggressive actions
	Power-gamers (Thue et al.)	Acquisition of special items and riches
	Tacticians (Thue et al.)	Creative thinking
	Storytellers (Thue et al.)	Interest in a complex plot
	Method Actors (Thue et al.)	Preference for dramatic actions
Bulatov	Awareness	Initial awareness
	Interest	Game purchase and installation
	Decision	Game acceptance or not
	Action	Playing and having interest in playing
Fairclough	Low Distress - Engagement	Comfortable zone (engagement aspect)
	Low Distress - Disengagement	Uncomfortable zone
	High Distress - Engagement	Stretch zone, comfortable zone (engagement aspect)
	High Distress - Disengagement	Uncomfortable zone
Brayshaw and Gordon	Competent	Achieving all game challenges with the best possible performance
	Autonomous	Customization of avatar, perception of self role and coherent behavior

To be continued

Model	**Profile**	**Characteristics**
	Relational	Social interaction, knowledge sharing, coaching, and social support
Smith	The Progress	Learning about game mechanisms
	The Fate	Internal probabilities in-game, and players' story controlled by destiny
	The Power	Disputes and contests; the wish to gain power
	The Identity	Positive social interactions as a member of a group and confirmation of self-identity
	The Imaginary	Imagination, creativity, and innovation aspects of players
	The Self	The playful pursuit of hobbies, assuming playing as an relaxing activity and escapism
	The Frivolous	The playful protest against social and cultural order of everyday life
Pine and Gilmore	Participation - Active	Interactions towards the environment
	Participation - Passive	Passive attention towards the environment
	Connection - Absortion	Learning without interaction
	Connection - Immersion	Physical or mental experience
Kellar et al.	Control (Kellar et al.)	Autonomy over tasks, social interactions, encouragement of innovation, proposition of goals, and guidance abilities
	Context	Interest in a complex plot, environmental beauty, and the receipt of feedbacks from actions
	Competence	The successful execution of complex tasks and proposition of strategies (identifying its effectiveness)
	Engagement	Personalization, acquisition of rewards, role-playing challenges, personal interpretation, and collaboration with others
Yee	Advancement	The desire to progress quickly, gain power, and accumulate riches and status
	Mechanics	Interest in analyzing game rules to optimize performance

To be continued

Model	Profile	Characteristics
	Competition (Yee)	The desire to compete against others (challenge)
	Socializing	Interest in helping and being in touch with others
	Relationship	The desire to establish long-term relationships
	Teamwork	Satisfaction of being part of an effort group
	Discovery (Yee)	The desire to discover things that others do not know about
	Role-Playing (Yee)	The creation of a persona with a background story and interactions with others to create an improvised story
	Customization	Customization of avatar (appearance)
	Escapism (Yee)	The avoidance of real life problems while playing
Malone and Lepper	Challenge (Malone and Lepper)	The pleasure to have the needed ability to solve a challenge (neither easy nor difficult)
	Curiosity	Attraction by the game environment (e.g., lights, sounds, and colors) and game story (desire to know everything)
	Control (Malone and Lepper)	The desire to feel important by controlling the environment and checking the results of the decisions made
	Fantasy (Malone and Lepper)	The desire to see images, physical, or social activities that do not correspond to daily life
	Cooperation and Competition	The desire to cooperate or compete with others
	Recognition	The receipt of approval (feedback) and recognition from others
Weiller	Checkpointer**	Progression in fantasy or realistic game story
	Solver**	Solving problems (achievement) in a satisfactory way
	Confident**	Obtainment of power and self-identification in an avatar
Koster	Enjoyment (Koster)	Enjoyment over discovering and mastering the game mechanisms
	Fruition	The pleasure of playing and doing activities linked to it; playing as a leisure activity

To be continued

Model	**Profile**	**Characteristics**
Lazzaro	Strong diversion (hard fun)	The accomplishment of a challenge
	Easy diversion (easy fun)	Aesthetic appreciation
	Altered States (relaxation)	Enjoyment based on excitement or relief (thrilling)
	Social Factor (amusement)	Social interactions
Colwell	Companionship	The desire to cooperate with others
	Prefer Friends	Preference of playing with friends
	Fun Challenge	Enjoyment entailed by challenges; playing as a funny and challenging activity
	Stress Relief	Tension release after completing a difficult challenge (i.e., the concept of catharsis)
VandenBerghe	Adventurer	Exploration of fantasy environments
	Investigator	Exploration of realistic environments
	Architect	The building of things in realistic environments
	Imagineer	The building of things in fantasy environments
	Masterer	Skilled and hard-working
	Perseverer	Not skilled and hard-working
	Dabbler	Impulsive and not skilled
	Talent	Impulsive and skilled
	Party Animal	Extrapolated behavior and social interactions
	Lone Wolf	Extrapolated behavior and absence of social interactions
	Hermit	Serene behavior and absence of social interactions
	Shepherd	Serene behavior and social interactions
	Sport	Teamwork in fantasy environments
	Knight	Player versus Player (PvP) battles in fantasy environments
	Killer (VandenBerghe)	Player versus Player (PvP) battles with complex fighting mechanisms
	Soldier	Teamwork with complex fighting mechanisms
Hunicke et al.	Sensation	A game as a sense of pleasure; playing as a pleasurable activity

To be continued

Model	Profile	Characteristics
	Fantasy (Hunicke et al.)	A game as make-believe
	Narrative	A game as an unfolding story
	Challenge (Hunicke et al.)	A game as an obstacle course
	Fellowship	A game as a social network
	Discovery (Hunicke et al.)	A game as uncharted territory
	Expression	A game as a "soap box" (i.e., a place to express an opinion or to be in evidence)
	Submission	A game as a mindless pastime; playing as a pastime activity
Cammarata et al.	Low trust	A high degree of shared capabilities
	Medium trust	A medium degree of shared capabilities
	High trust	A low degree of shared capabilities
Nacke et al.	Seeker	Explorer of the game-world
	Survivor	Preference for terror or fear challenges
	Daredevil	High risks associated with thrill
	Mastermind	The idea of the solver of problems (i.e., puzzles) and performance seeker (efficiency in decisions and strategies)
	Conqueror (Nacke et al.)	The overcoming of challenges (including other players or not)
	Socialiser (Nacke et al.)	Positive social interactions (e.g., talking, helping; trust relationships)
	Achiever (Nacke et al.)	Preference for accomplishing objectives
Caillois	Agon (Greek word)	It is the idea of challenges that entails in direct conflict or competition
	Aleais (Latin word)	It is the idea of games influenced by chance and randomness
	Mimicry (Greek word)	It regards the idea of role-playing, playacting, and dress-up
	Ilinxis (Greek word)	It refers to the thrilling sensation in gameplays (i.e., a visceral impact)
Yee and Bailenson	The proteus effect	A player changes his/her behavior according to the avatar's appearance

To be continued

Model	**Profile**	**Characteristics**
Harbord and Dempster	Anonymity**	Protection against embarrassment/shame provided by anonymity (escapism linked to personal-identification)
	Similar Appearance**	Players prefer to interact with players with an appearance near to their group or neutral appearances
Jerčić	Punter**	When a person is confident and tends to choose risky actions
	Cautious**	When a person realizes self-mistakes and tends to choose no risky actions
Reilly et al.	Explorer (Reilly et al.)**	When a player explores more than exploits
	Half-explorer half-exploiter**	When a player explores and exploits in the same degree
	Exploiter**	When a player exploits more than explores
Demetrovics et al.	Social	The pleasure in knowing people, being with others in a cooperative manner
	Escape (Demetrovics et al.)	To avoid problems of the real world (escapism)
	Competition (Demetrovics et al.)	The wish to compete and defeat others as a sense of achievement
	Coping	The improvement of mood derived from in-game challenges
	Skill Development	The wish to improve self skills, such as coordination and concentration
	Fantasy (Demetrovics et al.)	The enjoyment derived from assuming a new identity in a fantasy world linked to doing activities not possible in the real world
	Recreation	Playing as a relaxing and recreational activity
O'Brien and Toms	Point of Engagement	An interaction that affects positively the individual
	Sustained Engagement	Interactions with the object can maintain the individual affected positively
	Disengagement	Interactions with the object keep the individual affected negatively
	Extintion	The end of interactions between the object and the individual due to negative historical

To be continued

Model	**Profile**	**Characteristics**
	Reengagement	A situation when a negatively affected individual receives positive interactions and returns to a state of positively affected
Marczewski	Socializers (Marczewski)	When a person desires cooperation, competition, teamwork, social network, social pressure, social discovery, social enjoyment, and social status
	Free Spirits	When a person is live (i.e., enjoy easy fun), and creative and likes explorations (discovering things that others do not know), the autonomy to make choices, and customization
	Achievers (Marczewski)	When a person enjoys having hard fun, challenges, and boss battles, and wish to learn new skills, improve his/her progression, and finish quests
	Philanthropists	When a person enjoys having serious fun, identifying meanings, care-taking, receiving access, collecting, trading, gifting, and sharing (of items or knowledge)
	Players	When a person seeks rewards (e.g., points, EXP, prizes, badges, and money) and enjoys exhibitionism (e.g., top positions at leaderboards)
	Disruptors	When a person seeks changes regarding voting and anarchy, likes innovation, light touch, and anonymity
Sherry and Lucas	Competition/ Achievement	When a person wishes to be the best player, achieving the highest possible scores
	Challenge (Sherry and Lucas)	When a person is persistent to keep going in front to overcome the game challenges
	Social interaction (Sherry and Lucas)	When a person enjoys playing with friends and having social interactions
	Diversion/ Enjoyment	Playing as a pastime activity, and as a way to alleviate boredom

To be continued

Model	**Profile**	**Characteristics**
	Fantasy (Sherry and Lucas)	When a person wishes to do things that are not usually possible in real life
	Arousal	Playing as an exciting activity
Chou and Tsai	Entertainment	A game as a source of entertainment
	Escaping from people and routines	A game as an escapism environment
	Seeking information	A game as a source of information
	Escaping loneliness	A game as a tool to not feel alone
	Filling time	Playing as a filling-time activity
	Social device	A game as a tool to be with others
Kim and Ross	Sport lover**	When a person plays a game because it simulates a sport that he/she likes
Frostling-Henningsson	Cooperation	When a person performs tasks together with another
	Communication	When a person teaches, shares knowledge, or discuss personal problems
	Control (Frostling-Henningsson)	When a person feels that he/she has the needed ability to control a situation
	Escapism (Frostling-Henningsson)	A game as a place of refuge
	Hallucination of the real	When a person can do things that are usually not done in the real-world
Lin and Lin	Security (Lin and Lin)	When a person prefers safe gameplay, avoiding tough challenges regardless of the needed effort to allow that
	Fun and Enjoyment of Life	When a person has contact with novel content (e.g., new challenges or environments), improves performance, relieves stress, has fun, and interact toward others

To be continued

Model	**Profile**	**Characteristics**
	Warm Relationship With Others	When a person prefers to cooperate with others
	Sense of Belonging	When a person prefers to be part of a group
	Sense of Accomplishment	When a person enjoys winning related to fantasy
Jansz et al.	Fantasy (Jansz et al.)	When a person can do things in the virtual world that is usually not possible in the real-world
	Social interaction (Jansz et al.)	When a person enjoys having social interactions with friends
	Diversion (Jansz et al.)	When a person can do something uncommon, leaving away usual concerns
	Control (Jansz et al.)	When a person likes to have control of his/her avatars lives
	Challenge (Jansz et al.)	When a person wishes to improve his/her performance
	Enjoyment (Jansz et al.)	Playing as a fun activity
Olson	Hanging Out	When a person enjoys passing the time with friends
	The Joy of Competition	When a person enjoys competitions to beat others
	Teaching	When a person likes to teach others
	Making Friends	When a person wishes to make new friends
	Leadership	When a person likes motivating, persuading, and mediating a group of people
	Regulating Feelings	When a person plays to relax, cope with anger, and forget real-life problems
	Flow	When a person has pleasure in having the needed skill to solve a challenge in a goal-driven activity
	Challenge and Mastery	When a person wishes to master not trivial game mechanisms
	Expressing Creativity	When a person likes to create or customize content
	Different Identities	When a person likes to have different identities in-game

To be continued

Model	Profile	Characteristics
	Unreality	When a person enjoys doing things that cannot be done in the real-world
	Discovering	When a person wishes to learn new things
	Violence seekeer	When a person wishes to cope with violent situations to highlight self-status or learn how to overcome fearful situations
Eglesz et al.	Entertainer	Replaying as a relaxing, thrilling, interesting, and entertaining activity
	Explorer (Eglesz et al.)	When a person wishes to discover new things (e.g., new strategies or new ways to finish the game)
	Master (Eglesz et al.)	When a person wishes to improve performance by facing more difficult challenges
Griffiths et al.	Social features	When a player likes to have social contact, assist others, or be a part of a group
	Violent features	When a player enjoys PvP combats
	Playing alone features	When a player likes to play without social interactions
	Game progression features	When a player likes to progress in an endless game
	Avatar improving features	When a player likes to improve his/her avatar level and discover new strategies
Koo	Concentration	When a person likes to be aware of the surroundings and maintain a planned work
	Enjoyment (Koo)	Playing as an exciting, fun, and interesting activity
	Escape (Koo)	Playing as a way to refrain from feeling boredom
	Epistemic curiosity	Playing as a thinking and learning activity
	Social affiliation	When a person likes to be friendly with others, talkative, and a part of a group
Przybylski et al.	Harmonious Passion (HP)	Playing as a desire
	Obsessive Passion (OP)	Playing as a need
Wu et al.	Fairness	The game offers a fair trade regarding effort and reward. Cheating is not allowed

To be continued

Model	**Profile**	**Characteristics**
	Incentive	The game provides new challenges as well as rewards to the ones that keep playing
	Security (Wu et al.)	The game protects the players' data privacy as well as provide stable access to the game
	Stickness	When a person wishes to stay in the game longer than everyone
	Spatial Presence	When a person feels that the game is a part of his/her world
	Social presence	When a person likes to help others, receive support, expose self-values, be a member of a group
Hsu and Lu	Social Norms	When a person perceives that he/she should play a game to be in harmony with his/her group
	Critical Mass	When a person perceives that his/her group play a game frequently
Blacow	Power Gaming (Blacow)	When a person likes to have the power to defeat others in battles, having tendencies to perform treason, murder, and disturb others' gameplay
	Role-Playing (Blacow)	When a person likes to create an avatar that acts following strong self-beliefs
	Wargaming	When a person likes to think tactically to obtain better performance from the game mechanisms
	Story Telling (Blacow)	When a person wishes to understand the game tale beyond his/her avatar influences
Laws	Power Gamer (Laws)	When a person wishes to increase his/her power, strength, and riches and likes to identify opportunities with good benefits attached to low effort
	Butt- Kicker	When a player likes to prove his/her superiority to anyone who challenges him/her
	Tactician (Laws)	When a person likes to solve complex problems through rational thinking, measuring the efficiency of the decisions made. A person that is easily annoyed when there are others' opinions that go against his/her rational thought

To be continued

Model	Profile	Characteristics
	Specialist	When a person has a strong preference regarding a given feature or character, enjoying situations where such a feature or character has advantages towards others
	Method Actor (Laws)	When a person likes to express personal opinions, having a strong identification regarding his/her avatar. A person that likes to act when his/her group expects a reaction from him/her
	Storyteller (Laws)	When a person likes to unfold the game story, regardless of his/her avatar being a part of it or not
	Casual Gamer	When a person plays because of friends and not by internal motivations, this kind of person does not like to be forced to have a higher degree of participation than he/she feels comfortable
Dickey	Bounty	When a person likes to receive the reward associated with defeating an enemy or threat
	FedEx	When a person likes to receive the reward associated with delivering items
	Messenger	When a person likes to receive the reward associated with delivering a message
	Collection	When a person likes to receive the reward associated with collecting items (peacefully or not)
	Escort	When a person likes to receive the reward associated with escorting
	Goodwill	When a person likes to help others without wanting any reward
Haggis-Burridge	System immersion	When a person approves the game rules and mechanisms, being able to solve the provided challenges in a not trivial manner
	Spatial immersion	The sense of being present in the game-world
	Empathic/social immersion	The bond between a player and characters (NPC or human)

To be continued

Model	**Profile**	**Characteristics**
	Narrative/ sequential immersion	Progress in the game plot and environment exploration

Table 7 – Mapping between each player's model name and its reference

Model Name	**Reference**
Cook	(COOK, 2007)
Zhu et al.	(ZHU; LI; ZHAO, 2010)
Bartle	(BARTLE, 1996)
Bateman and Boon	(BATEMAN; BOON, 2006)
Thue et al.	(THUE et al., 2007)
Bulatov	(BULATOV, 2018)
Fairclough	(FAIRCLOUGH, 2008)
Brayshaw and Gordon	(BRAYSHAW; GORDON, 2016)
Smith	(SMITH, 1997)
Pine and Gilmore	(PINE; GILMORE, 1999)
Kellar et al.	(KELLAR; WATTERS; DUFFY, 2005)
Yee	(YEE, 2006)
Malone and Lepper	(MALONE; LEPPER, 1987)
Weiller	(WEILLER, 2015)
Koster	(KOSTER, 2005)
Lazzaro	(LAZZARO, 2004)
Colwell	(COLWELL, 2007)
VandenBerghe	(VANDENBERGHE, 2018)
Hunicke et al.	(HUNICKE; LEBLANC; ZUBEK, 2004)
Cammarata et al.	(CAMMARATA; KOSTER; Google's Advanced Technology and Projects (ATAP) group, 2018)
Nacke et al.	(NACKE; BATEMAN; MANDRYK, 2014)
Caillois	(CAILLOIS, 2001)
Yee and Bailenson	(YEE; BAILENSON, 2007)
Harbord and Dempster	(HARBORD; DEMPSTER, 2019)
Jerčić	(JERČIĆ, 2019)

To be continued

Model Name	Reference
Reilly et al.	(REILLY; ROOY; ANGUS, 2019)
Demetrovics et al.	(DEMETROVICS et al., 2011)
O'Brien and Toms	(O'BRIEN; TOMS, 2008)
Marczewski	(MARCZEWSKI, 2015)
Sherry and Lucas	(SHERRY et al., 2006)
Chou and Tsai	(CHOU; TSAI, 2007)
Kim and Ross	(KIM; ROSS, 2006)
Frostling-Henningsson	(FROSTLING-HENNINGSSON, 2009)
Lin and Lin	(LIN; LIN, 2011)
Jansz et al.	(JANSZ; AVIS; VOSMEER, 2010)
Olson	(OLSON, 2010)
Eglesz et al.	(EGLESZ et al., 2005)
Griffiths et al.	(GRIFFITHS; DAVIES; CHAPPELL, 2004)
Koo	(KOO, 2009)
Przybylski et al.	(PRZYBYLSKI et al., 2009)
Wu et al.	(WU; WANG; TSAI, 2010)
Hsu and Lu	(HSU; LU, 2004)
Blacow	(BLACOW, 1980)
Laws	(LAWS, 2002)
Dickey	(DICKEY, 2007)
Haggis-Burridge	(HAGGIS-BURRIDGE, 2020)

Based on the provided descriptions, most of the models portray players' interests, and few portray disengagement aspects associated with gaming. It is possible to identify different approaches, where some depict more general or abstract aspects, while others describe more specific ones. On the one hand, the abstract models usually use terms like fun, engagement, or enjoyment experiences, while the specific models provide more in-depth details of these terms. Another interesting observation was that the socialization aspect was present in 31 of the 46 models, suggesting that it is one of the main motivations to play games. In addition, Figure 6 shows the crescent number of players related models over the years considering the set of identified models. In special, we highlight the start of an increased rate in 2004, when more games started to adopt the GaaS policy following the success of games like World of Warcraft[1], demanding a more in-depth comprehension of the players' psychological aspects.

[1] Game's official site: <https://worldofwarcraft.com>, release dates: <https://en.wikipedia.org/wiki/World_of_Warcraft>

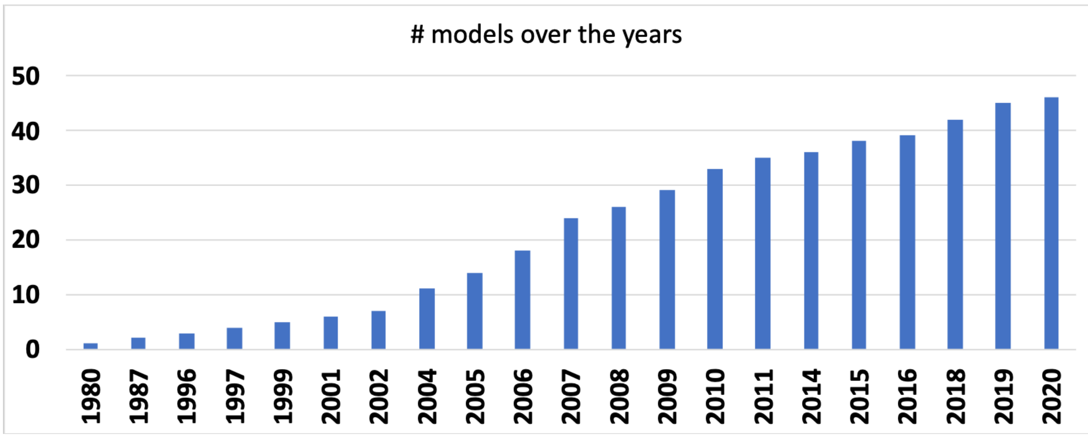


Figure 6 – The crescent number of players models over the years

2.6.3 Human-Being's Models

The human-being's models (HBMs) differ a little from the players' models, as they are not focused exclusively on games, portraying more general aspects. As the HBMs are more general than the players' ones, we assume that all player models are linked to them. Following the same procedure performed in the players' models case, all retrieved HBMs were read, summarized, and formatted in the proposed hierarchical structure (table format). Table 8 present all HBMs' descriptions. The reference map for each HBM can be found in Table 9.

Table 8 – Identified Human-Being's Models

Model	**Profile**	**Characteristics**
Myers et al.	Extroversion	Personal motivation comes from external factors.
	Introversion	Personal motivation comes from internal factors.
	Sensing	The acceptance of new information based on real facts.
	Intuition	The interpretation of new information based on abstract ideas.
	Thinking	The decision making guided by logic insights.
	Feeling	The decision making guided by the maintenance or increase of harmony.
	Judging	The preference in keeping the environment under control through planning.

To be continued

Model	**Profile**	**Characteristics**
	Perceiving	The preference in understanding and adapting to the environment.
Toprac and Abdel-Meguid	Anxiety	A psychological state that happens when something is likely to change, but a person does not know what or how.
	Suspense	A psychological state that happens when something will change.
	Fear	A psychological state that happens after a fearsome fact occurrence.
Frijda	Desire	The readiness of a person to approach, bring or have access to situations or things that give satisfaction.
	Happiness	A personal state of reasonable contentment.
	Interest	A personal interest in identifying new desirable things.
	Surprise	A personal reorientation after the occurrence of an unexpected fact (good or bad).
	Wonder	A personal reorientation that happens based on good facts.
	Sorrow	A psychological state wherein the mind passes to lesser perfection; when a person loses something or has less chances to have it.
Smith and Ellsworth	Pleasantness	An immediate and automatic interpretation of desires and possessions; the analyses of a new situation as desirable or not.
	Responsibility	A personal interpretation of actions in terms of social and personal standards; a final opinion about one's own actions.
	Certainty	A personal confidence that the environment will change.
	Attention	The others' perception toward one, disregarding one's opinion.
	Effort	A personal interpretation of a new situation that demands an action (an effort).
	Control	A personal control (or no control) over a situation.
Ekman	Anger	A person's antagonism toward an object, a person, or a group.

To be continued

Model	**Profile**	**Characteristics**
	Disgust	A response to something not desirable or offensive (e.g., social or physical unconformity).
	Fear	A person's request for help due to the anticipation of a threatening situation.
	Happiness	A person's well being and contentment.
	Sadness	A person's request for help as a response to a loss.
	Surprise	The response to an unexpected event.
Mehrabian	Pleasure	A polarity of an emotion as positive (high) or negative (low).
	Arousal	A psychological state degree ranging from alertness until calmness.
	Dominance	A personal control (or not) of the environment.
Huizinga	Diversion	The natural behavior of animals to explore and identify the possible interactions with the environment.
Deci and Ryan	Competence	The feeling of being capable, or not, to do some activity or assume a role.
	Autonomy	The freedom, or not, to make choices.
	Relatedness	The feeling of being connected, or not, to someone else.
Maslow	First Level	The attainment of physiological needs, such as eating, drinking, sleeping, etc.
	Second Level	The attainment of safety needs, such as personal and financial security, health, and well being.
	Third Level	The attainment of social needs, such as friendship, intimacy, and family.
	Fourth Level	The attainment of esteem needs, such as recognition, status, importance, and respect from others.
	Fifth Level	The attainment of personal dreams, such as mate acquisition, parenting, abilities usage, and goals achievement.
Murray	Materialism	The gain of possessions, the construction of something, the arrangement of objects, and the retention of objects.

To be continued

Model	**Profile**	**Characteristics**
	Power	The capacity of attacking or injuring, avoiding blame or punishment, revenging, and maintaining self-respect and pride in a high level.
	Affiliation	The capacity of accepting abasement, forming friendship, helping others, rejecting others, and seeking aid or protection.
	Achievement	The capacity of overcoming obstacles, resisting influence or coercion, avoiding pain, avoiding failure, being recognized, and claiming for attention.
	Information	The capacity of exploring, relating facts, and analyzing experiences.
	Sensual	The capacity of relaxing, enjoying sensuous expressions, and forming an erotic relationship.
Zillmann (Mood Theory)	Good Mood	A person's arrangement over the environment to maximize or maintain his/her pleasure.
	Bad Mood	A person's arrangement over the environment to diminish or alleviate his/her pain.
Zillmann (Affective Disposition Theory)	Positive Affect	When a person has empathy to another.
	Negative Affect	When a person has counter-empathy to another.
OCC	Joy	When a person is pleased due to the occurrence of a desirable event.
	Distress	When a person is upset due to the occurrence of an undesirable event.
	Happy-for	When a person is happy due to other's happiness.
	Pity	When a person is unhappy due to other's displeasure.
	Gloating	When a person is happy due to other's displeasure.
	Resentment	When a person is unhappy due to other's happiness.
	Hope	When a person is happy due to the prospect of a desirable event.
	Fear	When a person is upset due to the prospect of an undesirable event.

To be continued

Model	**Profile**	**Characteristics**
	Satisfaction	When a person is happy due to the occurrence of a likely event.
	Fears-confirmed	When a person is upset due to the occurrence of a likely event.
	Relief	When a person is happy due to the no occurrence of a likely event.
	Disappointment	When a person is upset due to the no occurrence of a likely event.
	Pride	When a person is admirably approving his/her own action.
	Shame	When a person is disapproving his/her own blameworthy action.
	Admiration	When a person is admirably approving someone else's action.
	Reproach	When a person is disapproving someone else's action.
	Gratification	When a person is admirably approving his/her own action and its consequences.
	Remorse	When a person is disapproving his/her own action and its consequences.
	Gratitude	When a person is admirably approving someone else's action and its consequences.
	Anger	When a person is disapproving someone else's action and its consequences.
	Love	When a person is liking appealing objects or persons, the more the object or person is known, the more the person likes it.
	Hate	When a person is disliking unappealing objects or persons.
Plutchik	Joy	When a person possesses or gains a valuable object.
	Trust	When a person has friends and is a member of a group.
	Anticipation	When a person is examining a new territory or situation.
	Surprise	When a person is trying to understand an unexpected event.
	Anger	When a person faces an obstacle (an enemy).

To be continued

Model	**Profile**	**Characteristics**
	Fear	When a person identifies a threatening situation.
	Disgust	When a person has an undesirable interaction with an object, a person, or a situation.
	Sadness	When a person loses a valuable object.
Nakamura and Csíkszentmihályi	Anxiety	When a person is pursuing a need and the challenge is too difficult according to his/her skills.
	Flow	When a person is pursuing a need and the challenge is neither too easy nor too difficult according to his/her skills.
	Boredom	When a person is pursuing a need and the challenge is too easy according to his/her skills.
Hunt	Positive Incongruity	When the context is more complex than the person's mental model of it; frustration.
	Low or no Incongruity	When the context complexity is as similar as the person's mental model of it; pleasure.
	Negative Incongruity	When the context is less complex than the person's mental model of it; boredom.
Goleman (The Big Eigth)	Fear	When a person identifies a dangerous or threatening situation that he/she does not feel able to deal with; a seeking for help.
	Anger	When a person's goal or an important thing is blocked or taken away.
	Sadness	When a person loses something (tangible or figurative); the loss of a wanted thing.
	Shame	When a person is afraid of being excluded from a group because of his/her actions or status.
	Disgust	When a person identifies gross, dangerous, or distasteful things.
	Jealousy	When a person has something valuable and it is in jeopardy of being taken away, or when a desirable thing is obtained by others; the control over the environment.

To be continued

Model	**Profile**	**Characteristics**
	Happiness	When a person gets what he/she wants, regarding physiological or psychological aspects, even though the way to get it was painful.
	Love	When a person has his/her relationship needs attained and feels valued and respected.
Goleman (Competences)	Self-awareness	When a person is able to: recognize his/her own emotions and their effects, know his/her strength and limits, and be sure about self-worth and capabilities.
	Self-regulation	When a person is able to: maintain a stable emotional state with fair and honest behavior, be responsible for performances, and be receptive to beneficial changes.
	Self-motivation	When a person is able to: improve own performance, be aligned with the objectives of a group, be persistent, and prompt to act.
	Empathy	When a person is able to: anticipate, recognize, and meet others' needs, as well as improve others' skills, strengthen relationships and foster progress opportunities.
	Social Skills	When a person is able to: persuade, motivate, guide, propose changes, resolve disagreements, and boost others.
Goldberg	Openness to Experience	It regards aspects of internal knowledge representation, environment perception, imagination, curiosity, and creativity.
	Conscientiousness	It regards aspects of cautiousness, persevering, responsibility, carefulness, and hardworking.
	Extraversion	It regards aspects of socialization, sincerity, enthusiasm, and the seeking for positive emotion.
	Agreeableness	It regards friendship, altruism, cordiality, and confidence.
	Neuroticism	It regards aspects of insecurity, auto-control, anxiety, irritability, fright, and uneasiness.
Eysenck	Phlegmatic	When a person is passive, thoughtful, reliable, and peaceful; a relation between low neuroticism and introversion.

To be continued

Model	**Profile**	**Characteristics**
	Melancholic	When a person is quiet, pessimist, unsociable, sober, rigid, moody, anxious, and reserved; a relation between high neuroticism and introversion.
	Sanguine	When a person is sociable, relaxed, serene, leaderly, optimistic, and reactive; a relation between low neuroticism and extraversion.
	Choleric	When a person is impulsive, instable, uneasy, aggressive, and optmistic; a relation between high neuroticism and extraversion.
Northrop	Abasement	When a person is humble, resigned, servent, discreet, self-belittling, and accepts punishment even when not deserved.
	Achievement	When a person is hard-working, resourceful, achiever, competitor, and perfectionist.
	Affiliation	When a person appreciates the presence of friends and familiars and likes to maintain and create new friendships.
	Aggression	When a person appreciates combats and discussions, is moody, and does not care about others, hurting them to achieve own objectives if it is needed.
	Autonomy	When a person appreciates the freedom to follow own way and rules, breaking restraints in the process if needed; lonely.
	Change	When a person appreciates new experiences, is adaptable, avoids routine, and frequently changes his/her opinions or values depending on the circumstances.
	Cognitive Structure	When a person does not like uncertainty (incomplete information) and prefers to make decisions based on real facts, rather than guesses or probabilities.
	Defendence	When a person promptly protects him/herself from threats, harms, or criticism, taking an offensive position if needed.

To be continued

Model	**Profile**	**Characteristics**
	Dominance	When a person wishes to control and influence the environment through impositions.
	Endurance	When a person strongly perseveres to accomplish an objective, following a strict routine and working many hours to overcome whatever is the obstacle, hardly giving up of it.
	Exhibition	When a person wants to be in evidence toward others, being dramatic, if needed, to achieve it.
	Harmavoidance	When a person prefers safe activities instead of exciting ones to avoid harm, seeking for support if needed.
	Impulsivity	When a person reacts without deliberation, presenting an unstable emotional state.
	Nurturance	When a person supports others always as possible, giving comfort and doing favors.
	Order	When a person likes to keep the surroundings organized, as well as his/her personal belongings.
	Play	When a person likes to spend time in cheerful activities, enjoying funny moments with a light-heart and no concerns.
	Sentience	When a person associates environment characteristics, such as smells, sounds, sights, tastes, and textures to important things of life.
	Social Recognition	When a person wishes to receive recognition from others and be held in high esteem reputation in a society.
	Succorance	When a person frequently needs to receive the support and care from others, such as love, sympathy, protection, advice, and reassurance to avoid insecurity.
	Understanding	When a person desires to acquire knowledge from many different areas to satisfy his/her curiosity, verifying generalizations and synthesis of ideas through logical thought.

To be continued

Model	**Profile**	**Characteristics**
	Infrequency	When a person reacts in a careless way, presenting implausible or pseudo-random behavior due to poor comprehension, confusion, or passive non-compliance.
	Desirability	When a person wishes to foment a favorable self-image toward others (consciously or not), describing him/herself in desirable terms regardless if they are accurate or not.

Table 9 – Mapping between each HBM name and its references

Model Name	**Reference**
Myers et al.	(MYERS et al., 1998)
Toprac and Abdel-Meguid	(TOPRAC; ABDEL-MEGUID, 2011)
Frijda	(FRIJDA, 1986)
Smith and Ellsworth	(SMITH; ELLSWORTH, 1985)
Ekman	(EKMAN, 1992)
Mehrabian	(MEHRABIAN, 1980), (MEHRABIAN, 1995), (MEHRABIAN, 1996), (RUSSELL; MEHRABIAN, 1977)
Huizinga	(HUIZINGA, 2014)
Deci and Ryan	(DECI; RYAN, 1985), (DECI; RYAN, 1995)
Maslow	(MASLOW, 1968)
Murray	(MURRAY, 1938)
Zillmann (Mood Theory)	(ZILLMANN, 1988), (ZILLMANN, 2015)
Zillmann (Affective Disposition Theory)	(ZILLMANN, 1995), (ZILLMANN, 1996)
OCC	(ORTONY; CLORE; COLLINS, 1990)
Plutchik	(PLUTCHIK, 1980)
Nakamura and Csíkszentmihályi	(NAKAMURA; CSIKSZENTMIHALYI, 2009), (CZISIKSZENTMIHALYI, 1990)
Hunt	(HUNT, 1963)
Goleman (The Big Eigth)	(GOLEMAN, 1995)
Goleman (Competences)	(GOLEMAN, 1998)

To be continued

Model Name	**Reference**
Goldberg	(GOLDBERG, 1990)
Eysenck	(EYSENCK, 1967), (EYSENCK, 1973)
Northrop	(NORTHROP, 1974), (NORTHROP, 1984)

Unlike the players' models, where preferences, stages, and motivational aspects were portrayed, the HBMs depicted different aspects not limited to contexts, such as emotions, personality traits, human needs, and personal competencies. Even though each model focused on a core aspect, we could identify correspondences between different approaches. Therefore, we suggest that the human behavior can be portrayed as a graph (see Figure 7), where each node is a model and is connected to another somehow. For example, a personality trait depicts how a person behaves, entailing in defining what human needs he/she will pursue, which will generate emotions associated with the attainment or not of such needs. Also, the historical occurrences of positive and negative emotions entail in positive, neutral, or negative sentiments regarding a subject. This aspect will be deeply explored in Chapter 5.

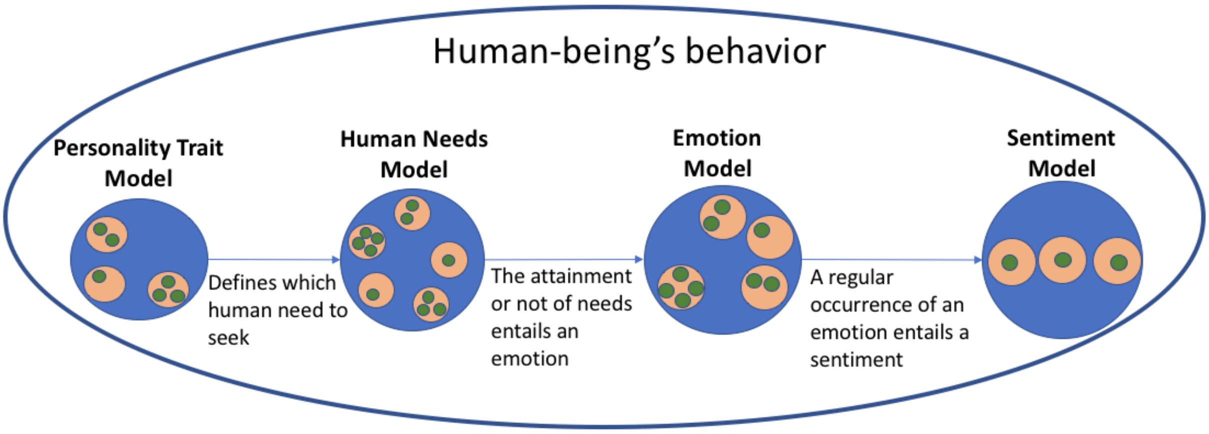


Figure 7 – Proposed human-being graphical behavior

As occurred to the players' models, the social aspect is also one of the main aspects presented by HBMs; from the total of 21 models, 14 consider it. In conclusion, we can suggest that social factor is one essence of humans' behavior. In addition, Figure 8 shows the crescent number of HBMs propositions and enhancements over the years, considering the set of identified models. It is possible to notice that the HBMs range of years is wider than the players' models; what is expected as players models encompasses a more current context.

2.7 Related Works

The psychological models portrayed in the Retrieved Works Section 2.6 were obtained from works that used or cited them somehow, and between them, only a few

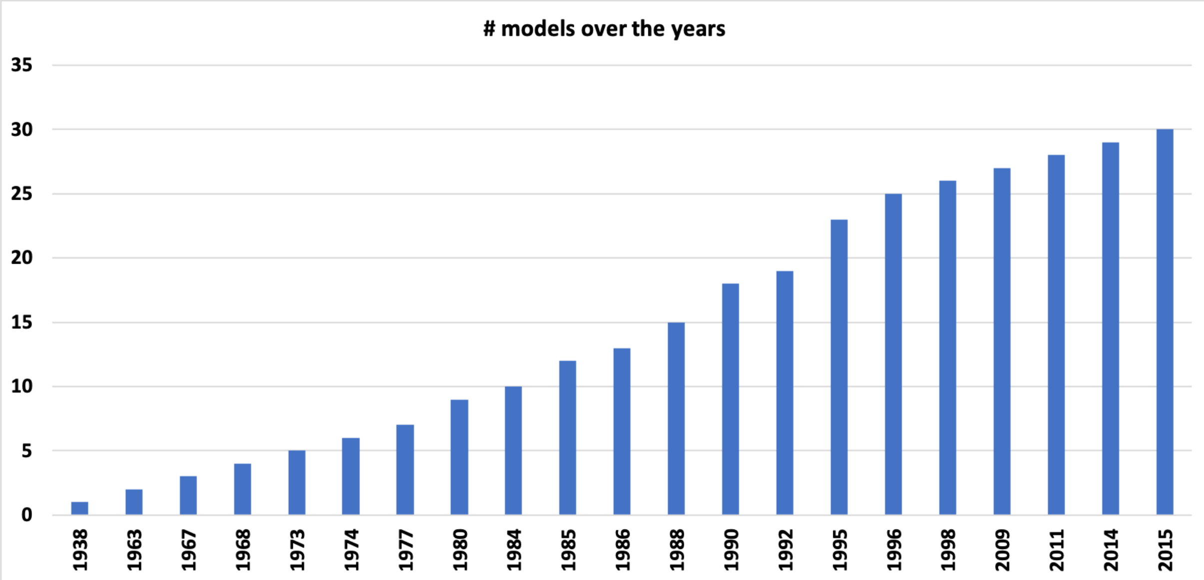


Figure 8 – The crescent number of propositions and enhancements of HBMs over the years

ones were similar to our objective (the identification of players' psychological profiles in usage data based on the findings of psychological models). In particular, we describe the works that focused on the following topics:

1. Game Analytics approaches that adopted psychological models.
2. Game development support.
3. Personality traits identification in usage data.
4. Emotion simulation in NPCs (non-player characters). The idea of player simulation.
5. Association between human needs and actions in-game.

The identified Game Analytics approaches regarded the works of (BONOMETTI et al., 2019; JEON et al., 2017; MAKANTASIS; LIAPIS; YANNAKAKIS, 2019). The first work adopted the O'Brien and Toms model (O'BRIEN; TOMS, 2008), whereas the second used notions from the Yee (YEE, 2006), and Goldberg (GOLDBERG, 1990) models, and the third ideas of the players' arousal model of Mehrabian (MEHRABIAN, 1980; MEHRABIAN, 1995; MEHRABIAN, 1996; RUSSELL; MEHRABIAN, 1977). All of them proposed identifying psychological features, where some of them applied these features to the churn prediction problem, while others to the measurement of players' pleasure in play. Another approach regarded the work of (CAMILLERI; YANNAKAKIS; LIAPIS, 2017). This study proposed a prototype of a general model of affect to games based on the Russel model (the base of the Mehrabian model). Despite its interesting proposition that encompasses multiple games, it is limited due to the inherit mid-term representation of the arousal aspect, where short, and long-term psychological aspects (such as human needs attainments and personalities nuances, respectively) are not considered (reducing the proposed notion of "general"). Even with the interesting propositions linked to the adoption of psychological models in the Game Analytics field, most of the authors adopted only one model and not systematically justified this adoption by considering other models, limiting the extension of their findings.

Also linked to the Game Analytics, an exciting work (ZHAO et al., 2020) proposed the identification of players' preferences through the analysis of pieces of players' actions sequences. This idea is similar to one present in the proposed method introduced in Chapter 5.

Moving to the game development support, the work of (SNODGRASS; MOHADDESI; HARTEVELD, 2019) proposed the PEAS framework, which links psychological models to game components development. Even though it can consider multiple psychological models simultaneously, no procedure is present in the framework regarding the identification of a general model. Such identification is the aim of the Unification Explorer Framework (UEF) presented in Chapter 4.

The automatic identification of personality traits (long-term aspects) in usage data was the research subject from Giel van Lankveld on his thesis (LANKVELD, 2013). He developed games that presented choices to players linked to answers of questionnaires used to identify personality traits (regarding The Big Five or FFM model (GOLDBERG, 1990)). Therefore, while a player plays, he/she is answering the questionnaire (unbeknownst of doing it). It is interesting to highlight that one of his future works is exactly what we are proposing (specially, in Chapter 7), he stated "a careful validation of game behavior for a psychological construct is an essential task in game research that has not yet been properly addressed".

The simulation of emotions (short-term aspects) in NPCs was the subject of study for Popescu, Broekens, and Someren in (POPESCU; BROEKENS; SOMEREN, 2014). They proposed the named "Gamygdala" (in analogy to the amygdala, the part of the brain responsible for human emotions) which is a game engine that helps game developers add emotions into NPCs. They proposed a way to simulate 16 of the total 22 emotions of the OCC model (ORTONY; CLORE; COLLINS, 1990) portraying internal emotions and social emotions. Internal emotions regard only aspects of the self (the NPC), and social ones regard the relation with other NPCs and the players. The three key concepts used to simulate the internal emotions were (1) the definition of a goal (to wish or avoid something; desirability), (2) the likelihood of achieving such goal, and (3) the final result (if the NPC achieved it or not). Examples of rules to simulate emotions are presented in Table 10. It is possible to divide those emotions into three categories regarding the expectancy (Hope and Fear), certainty (Joy and Distress), and "denouement" (Satisfaction, Fears-confirmed, Disappointment, and Relief). The social emotions were simulated based on two key concepts, (1) the changes of the likelihood of an event caused by others or toward others (NPCs or players; the causer and the affected) and (2) the appreciation or disdain of the self towards the causer or the affected (examples are presented in Table 11). It is also possible to divide the social emotions into two categories, wherein one category the self considers who is the causer or the affected (Pity, and Gloating) and in another where the causer is not considered (Anger, Guilt, and Gratitude). One example of an internal emotion occurrence in an RPG is when a hero (NPC) wants to save a princess (goal), and it discovers a superweapon that increases the likelihood of saving her, entailing the Hope emotion. Linked to it, one example of a social emotion could be Gratitude if another NPC or player gave this superweapon to the hero.

Table 10 – Rules to simulate internal emotions, adapted from (POPESCU; BROEKENS; SOMEREN, 2014)

Emotion	**Eliciting Condition**
Hope	when there is something desirable with an increased likelihood or there is something not desirable with a decreased likelihood
Fear	when there is something not desirable with an increased likelihood or there is something desirable with a decreased likelihood
Joy	the certainty that something good will happen or the certainty that something not good will not happen
Distress	the certainty that something not good will happen or the certainty that something good will not happen
Satisfaction	when a desirable thing with an increased likelihood happened or a not desirable thing with a decreased likelihood not happened
Fears-confirmed	when a not desirable thing with an increased likelihood happened or a desirable thing with a decreased likelihood not happened
Disappointment	when a desirable thing with an increased likelihood not happened or a not desirable thing with a decreased likelihood happened
Relief	when a not desirable thing with an increased likelihood not happened or a desirable thing with a decreased likelihood happened

Table 11 – Rules to simulate social emotions, adapted from (POPESCU; BROEKENS; SOMEREN, 2014)

Emotion	**Eliciting Condition**
Anger	when the likelihood of something not desired increased or the likelihood of something desired decreased due to an action of another
Guilt	when the likelihood of something not desired by other increased or the likelihood of something desired by other decreased due to a self-action
Gratitude	when the likelihood of something desired increased or the likelihood of something not desired decreased due to an action of another
Pity	when an undesirable event happens to a liked NPC
Gloating	when an undesirable event happens to a disliked NPC

The last related work regards the association done by Bostan (BOSTAN, 2009) to link Murray's model of human needs (MURRAY, 1938) (short-term aspects) to RPGs. Each group of human needs was linked to several actions done in-game such as follows:

- Materialism

  Objects acquisition: it can be done through looting, quests rewards, stealing, opening hidden treasures, or buying.

  Objects order: the in-game inventory allows the organization of the obtained items.

  Possessions retention: a player can put items into safe-deposit boxes to avoid getting them lost.

  Construction of objects: a player can construct new objects or reforge them through interactions with forges or NPCs.

- Power

  Aggression: a player must attack and kill enemies to progress in the game story. However, in some games, a choice is offered to the player, for example, to kill or not an enemy, entailing changes in the game plot sequence.

  Blame avoidance and Defendance: in RPGs, it is usually possible to choose between "good" and "bad" options. Each decision can bring desirable and undesirable consequences to the player, who must be ready to justify them to avoid punishment.

  Counter action: when a player overcomes a defeat or a failure by restriving and retaliating (for honor or to avoid humiliation).

  Deference: RPGs can present a guild system, where the player can support and work with other players or NPCs (obeying a hierarchy and a division of tasks).

  Dominance: also considering the guild system, a player with a high position in the hierarchy may influence or control NPCs or players of lower positions.

- Affiliation

  Abasement: when a player accepts the punishment, apologizes or reconciles toward affected NPCs or players.

  Affiliation: when a player forms friendships to others (e.g., joining a guild).

  Nurturance: when a player helps others through donations or joining in others' parties to complete their quests.

  Rejection: when a player declines a new member or rejects to enter in a guild. The affiliation need is both attained and impaired in such situations (MURRAY, 1938).

  Succorance: when a game challenge is too difficult and a player must seek help to accomplish it.

- Achievement

  Achievement: the progress in-game, which is usually linked to acquiring experience points and new skills.

  Autonomy: as a player becomes stronger, he/she may depend less on others entailing in more significant autonomy (freedom) to make choices due to a lack of deliberations from others.

  Harm avoidance: when a player is able to stay "alive" after a challenge, usually using potions, spells, or plants to keep him/her healthy.

  Inf avoidance: when a player refrains from a very difficult challenge, avoiding shame and humiliation.

  Recognition: some RPGs have a reputation/fame system (points), which allows the player to be recognized by his/her accomplishment of difficult challenges.

  Exhibition: when a player attracts attention and thrills others.

- Information

  Cognizance: when a player is able to satisfy his/her curiosity about the game story through questions, observations, listenings, readings, and examinations.

  Exposition: when a player is able to explain, teach, or exchange information with others.

  Understanding: when a player is able to analyze the game environment and understand what is his/her role in it.

- Sensual

  Play: when a player can relax, amuse, play, laugh, joke, and be merry.

  Sentience: when a player seeks and enjoys sensuous expressions, usually towards a possible relationship.

  Sex: when a player has sensual inter-course in-game (an erotic relationship).

It is interesting to highlight that the relationship between human needs and in-game possibilities was done theoretically, not being possible to assess this linkage, as different people may have different interpretations. In Chapter 5, the works of (POPESCU; BROEKENS; SOMEREN, 2014; BOSTAN, 2009) are revisited.

2.8 Discussion and Analysis

This section is focused on answering the four research questions associated with this SLR based on the retrieved works.

2.8.1 Answer to RQ1

Research question 1 states the following "*What is a psychological aspect?*". We answer this question with the information described at the beginning of Chapter 2. A psychological aspect can be understood in more specific terms such as affect, emotion, sentiment, personality, personality traits, competence, and human needs.

2.8.2 Answer to RQ2

The research question 2, "*What is a psychological profile?*", can be answered with the proposed architecture presented in Figure 5. In our concept, a psychological profile is a set of characteristics (i.e., psychological aspects) that can describe one's behavior or way of being. In addition to it, psychological profiles are grouped inside psychological models, where different profiles of the same model have different characteristics. From a Game Analytics perspective, the idea of a psychological profile is the same, as it regards the set of psychological features, differentiating only how they are obtained (from usage data, instead of questionnaires or observations).

2.8.3 Answer to RQ3

The answer to "*What are the psychological models applied to games?*" can be summarized by Tables 6 and 8. The models portrayed there are based on players' behavior or were linked to games somehow.

2.8.4 Partial answer to RQ8

The initial answer for the research question 8, "*Are all models applicable to all game genres?*" is no, they are not. Using the model of Toprac and Abdel-Meguid as an example (TOPRAC; ABDEL-MEGUID, 2011), their profiles can be applied to horror games, so they may not fit a board game like Chess or an arcade game like Pac-Man, for example.

Another point can be discussed in the context of human needs of Murray (MURRAY, 1938), as different game genres satisfy different needs. For example, strategy games tend to satisfy the Materialism group, while social games tend to satisfy Affiliation, and RPGs to satisfy Power, Materialism, and Affiliation (BOSTAN, 2009). As we can see, RPGs usually attain more than one need as this kind of game offers virtual environments analogous to the real world. Finally, we can conclude that game genres may be grouped according to which types of human needs they attain.

2.8.5 Partial answer to RQ9

Based on the retrieved works, it was possible to initially identify three advantages entailed by using psychological models to answer the research question "*What are the advantages and disadvantages of using psychological models?*" regarding (1) new game contents, (2) believable NPCs, and (3) Game Analytics predictions.

When a new game version is in the designing process, identifying the players' motivations could guide the development of a new game content that should please players more (GUARDASCIONE, 2018; HARRISON et al., 2015). For example, if a game producer knows which kind of human needs his/her users most seek, a new game content focusing on those aspects may be successful.

NPCs' modeling process presents many challenges, being one of them the so-called "believability" (KERSJES; SPRONCK, 2016) (also known as Player Simulation). When a game has NPCs with emotions and personality, it gives uncertainty about the NPCs' behavior. For example, tense and nervous individuals make irrational and unpredictable choices (POPESCU; BROEKENS; SOMEREN, 2014; ROSENTHAL; CONGDON, 2012), giving the idea of "illusion of life" (BATES et al., 1994). In addition to it, there are some studies which state that players have more fun when playing against other players (KELLAR; WATTERS; DUFFY, 2005; KERSJES; SPRONCK, 2016) because they can socialize (SWEETSER et al., 2003). According to Kersjes and Spronck (KERSJES; SPRONCK, 2016), there are three main aspects to simulate emotions with credibility: (1) the identification of specific situations where they happen, (2) the ideal representation of emotions, and (3) the definition of a proper response. Therefore, in an ideal scenario, a believable NPC may give the same enjoyment as a human player does, improving the game experience. Frameworks to add emotions to NPCs can be found at (POPESCU; BROEKENS; SOMEREN, 2014; JOHANSSON; DELL'ACQUA, 2012a; JOHANSSON; DELL'ACQUA, 2012b; PANUMATE; MIYAKE; IIDA, 2016; HOLMGÅRD et al., 2014) and an application of believable bots (NPCs) into a crisis simulation system can be found at (LOIZOU et al., 2012). It is interesting to highlight that the application of believable bots to crisis simulation systems represents a crucial aspect, as in such cases of panic, people may behave based more on emotions than on logic.

The use of psychological models in the Game Analytics field has shown some improvements in identifying risk situations. The works of Kummer et al. (KUMMER; NIEVOLA; PARAISO, 2017b; KUMMER; NIEVOLA; PARAISO, 2018b) explored the engagement aspects of players based on usage data. They could identify inside the number of active players different degrees of commitment to the game, which allowed identifying psychological profiles, such as the Cook's ones (COOK, 2007). In those cases, the use of psychological models could highlight risk situations that were not possible to be identified until then, showing to game producers that some situations considered as good actually are dangerous. Therefore, identifying more descriptions of psychological models based on usage data may improve the quality of the risk management done by game producers.

Although there are advantages, we can point out that a disadvantage (challenge) is identifying an ideal model to be applied to a specific situation. For example, there are many models of emotions (TOPRAC; ABDEL-MEGUID, 2011; FRIJDA, 1986; EKMAN, 1992; ORTONY; CLORE; COLLINS, 1990; PLUTCHIK, 1980; GOLEMAN, 1995), but knowing which one is the best for a given situation is not a trivial task. In this work, we propose a solution to this problem by unifying different psychological models to identify a single one that carries all models' points of view. This suggestion is explored in the Chapter 4, where the Unification Explorer Framework is presented and applied to the sets of players' models and HBMs identified in the current Chapter.

2.9 Chapter Conclusions

Even though it is possible that this research does not portray all the psychological models that exist, being it a possible introduction, all the psychological models portrayed were linked to games somehow, so we consider it as a good starting point because their applications have previous experiments in the game field. Another interesting aspect of psychological models regards their conception. Where in one hand, most of the players' models are not based on HBMs, being them based only on players aspects, while on the other hand, models such as the VandenBerghe (VANDENBERGHE, 2018) and Bateman and Boon (BATEMAN; BOON, 2006) are totally based on more general human aspects. Some HBMs were also proposed based on others HBMs and no HBMs depicted in this review were based on players' models. In conclusion, we consider that HBMs have the potential to improve the understandings about players' models (even though some of them were not based on HBMs), justifying in that way the snowball process depicted in subsection 2.5.

The analysis of models' relationship seems to be promising, as more general conclusions can be inferred from them. One example is the behavioral graph idea (depicted in subsection 2.6.3), where even though each model may depict detailed aspects, connections between those aspects can be identified, allowing in that way the identification of more general behaviors.

Besides the knowledge provided by this Chapter, the next Chapter presents supplementary pieces of information required to understand further analysis, propositions, and discussions of this thesis. After it, in Chapter 4, the Unification Explorer Framework is presented and applied. Later, in Chapter 5, the UEF findings are used in the proposed method to identify players' psychological profiles on usage data.

3 Theoretical Background

This Chapter describes aspects regarding the MMORPG genre (Section 3.1), usage data (Section 3.2), the "Game Path" concept (Section 3.3), risk situations (Section 3.4), Game Refinement Theory (Section 3.5), Knowledge Discovery in Databases (KDD) (Section 3.6), the generation of the Commitment metric (Section 3.7), and Concept Lattices (Section 3.8).

3.1 MMORPG Genre

Role-playing games have their starting point outside the digital area. In the beginning, RPGs were played using only pens, papers, dice, and imagination (LAWS, 2002). In this old fashioned way, among the players, there is a special role called game master. The game master describes the game environment to the players, defines and applies rules to guide the possible interactions, and presents the challenges to overcome. Moreover, there are RPG books that help game masters, giving some predefined settings[1]. Usually, in an RPG, a player is represented by an avatar in the fictitious game world[2]. This avatar has a name, attributes (like strength, dexterity, and intelligence) that grow during game-play, pieces of equipment, an inventory, and an objective (e.g., obtaining a treasure; a quest).

During the course of a game, players encounter opponents with whom they have to fight, getting new items and experience in case of winning. The accumulation of experience points is used to define when a player "levels up", i.e., when a player improves his/her attributes. When a player levels up, he/she usually can choose which attribute to improve. For example, a warrior may choose to improve strength, while a magician may choose to improve intelligence. The players' attributes can be used to define when a player can use an item, apply a spell, or set his/her defensive and offensive powers; therefore, an avatar performance is based on its attributes and equipment. During the game story, players face many different quests, which may be attached to the main quest or not (players can choose to do or not such quests, but in some cases, the accomplishment of some quests are needed to develop the game story).

---

1 A famous classic RPG is the so-called Dungeons & Dragons from Gary Gygax and Dave Arneson, published in 1974. For more information please visit this website <https://en.wikipedia.org/wiki/Dungeons_%26_Dragons>

2 In the RPG case, the terms of player and avatar can be used as synonyms because it is common to a player control only one avatar; therefore everything associated with the avatar refers also to its player, and vice-versa. However, the same cannot be assumed to all game genres because there are cases where a single player controls many avatars simultaneously (e.g., in RTS games).

The game mechanisms are firmly attached to the randomness use of dice. For example, to a player successfully hit an opponent, a value greater than two must be obtained; to obtain a good price at a shop, a value greater than five must be obtained; or to reforge a weapon, a value greater than 15 must be obtained (i.e., a threshold value). There are dice with many sides in RPGs, such as six, eight, twelve, or twenty. A player's attributes may change the threshold set to a die for a specific action. Players may choose to cooperate or not with other players. As a final remark, classic RPGs are usually played in a table containing a map, dice, and a paper (sheet) for each player, which describes his/her items, attributes, skills, equipment, and inventory.

In the modern fashioned way (i.e., "the digital era"), the game master role is implemented by software, which implies that all the environment, rules, story, and interaction options are predefined, or in other words, they are fixed (in classic RPGs the game master can expand the game story and environment as much as he/she wishes during the gameplay). Moreover, the gameplay takes place in a digital environment (provided by a computer or console) in which a player has the same characteristics of classic RPGs, such as an avatar with attributes, items, equipment, and a mission.

An MMORPG is described then as a digital RPG played in an online environment with a massive number of players. A common situation that happens during the creation of an avatar in MMORPGs is the choice of the avatar's nation (or faction). Usually, this nation is at war with another, so players will eventually fight each other, the so-called Player versus Player (PvP). In this kind of game genre, when a player finishes the game story, he/she may continue playing to engage in PvP battles and improve his/her avatar's performance (i.e., more experience points, more attribute points, and better equipment). In addition to it, social interactions are very present in this kind of game. Players can create or join guilds, make friends, start conflicts, and even get married (in-game). Moreover, players can build houses and modify the game environment (not only using it). In conclusion, the MMORPG genre is a game genre known to provide a wide range of possibilities in social, competitive, and creation aspects (BOSTAN, 2009). Those characteristics will be deeply explored in Chapter 5, during the method proposition.

3.2 Usage Data

As previously introduced in Chapter 1, usage data are the data generated while players are playing digital games. Such kind of data is used by game producers to try to identify opportunities to act or risk situations to fix (KUMMER; NIEVOLA; PARAISO, 2017a). Moreover, usage metrics can be computed based on this data (e.g., Equations 1.1, 1.2, 1.3, 1.4, 1.5, among others).

Usage data portray players' actions (or events that affected the players), status, or final results, also having different granularities. Tables 12, 13, and 14 illustrate some examples of usage data with low, average, and high granularity (this measure is a pragmatic consideration, only to show that differences in granularity exist), portraying respectively the games World of Warcraft[3] (MMORPG), Blade&Soul[4] (MMORPG), and StarCraft II[5] (real-time strategy; RTS).

Table 12 – An adapted sample from the WOWAH (World of Warcraft Avatar History) dataset (LEE et al., 2011)

Timestamp	Avatar ID	Guild	Level	Race	Class	Zone
12/31/05 23:59:46	1	none	9	Orc	Shaman	Durotar
12/31/05 23:59:57	9	1	27	Orc	Hunter	Stonetalon Mountains

In Table 12, Timestamp is the exact data collect time, Avatar ID is the avatar's unique identification, Guild is the avatar's guild unique identification, Level is the avatar's current level, Race is the avatar's race, Class is the avatar's specialty, and Zone is the avatar's current place. The WOWAH dataset has 91,065 avatars. In Table 12, only two instances regarding two avatars are described, but it is essential to highlight that there is no limitation to the number of instances each avatar can have, regardless of game and data granularity.

Table 13 – An adapted sample from the Blade&Soul dataset (LEE et al., 2018)

Timestamp	Avatar ID	Level	Race	Job	Faction	Action
2016-04-20 23:16:26.497	00C172F0	50	4	9	2	Get Item
2016-04-11 18:46:08.605	0D45FEED	50	2	10	2	Get Money

Regarding Table 13, the columns regarding Timestamp, Avatar ID, Level, and Race have the same meaning as those depicted in Table 12 (just the Race value is an ID in this case). The Job column refers to the avatar's activity (a role ID), the Faction one the avatar's nation (as explained in Section 3.1), and the Action one the action done by the avatar in the given timestamp. This dataset has 10,000 avatars.

Table 14 – A match final result from StarCraft II (URIARTE, 2017)

Date	Team ID	Resources	Units	Structures	Overview
2010-02-18	1	30,650	59,425	28,000	123,725
2010-02-18	2	31,487	30,300	8,550	77,437

In view of Table 14, Date is the date when the match happened, Team ID is the team identification, Resources is the score associated with the gathering of materials,

[3] For more information about World of Warcraft please visit the following website: <https://worldofwarcraft.com/en-us/>

[4] For more information about Blade&Soul please visit the following website: <https://www.bladeandsoul.com/en/game/theater-of-mystery/>

[5] For more information about StarCraft II please visit the following website: <https://starcraft2.com/en-us/>

Units is the score associated with the production of combat units, Structures is the score regarding buildings, and Overview is the final score. This dataset has approximately 20,680 matches.

It is possible to identify a difference in granularity between the three datasets. The final result of a StarCraft II match presents only summarized information, while the WOWAH dataset presents a set of players' status over time, and the Blade&Soul dataset presents every action done or events received by players (together with the players' status). Therefore, we consider the StarCraft II, WOWAH, and Blade&Soul datasets as having low, average, and high granularity, respectively. In addition to it, common characteristics of those datasets are time, players' identification, and players' status.

A granularity degree implies the size of a dataset. An instance in a dataset can be understood as a player's status (or action/event) in a given time. The same player can have many instances in the same dataset, allowing in that way the tracking of individual behavior (the essence of Game Analytics, and also a possible approach of game producers). Table 15 illustrates the dataset size for the three games, as mentioned earlier, according to their granularity degree and collection period. As we can see, the higher the granularity degree, the higher the number of instances per player.

Table 15 – A relation between granularity degree and number of instances

Dataset	**Period of Collect**	**Granularity Degree**	**Number of Instances**	**Number of Distinct Players**	**Mean of Instances per player**
StarCraft II	Unknown	Low	41,360	Unkown	Unkown
WOWAH	3 years	Average	36,513,647	91,065	400.96
Blade&Soul	24 weeks	High	579,560,548	10,000	57,956

3.3 Game Path Concept

As depicted in Section 3.1, MMORPGs have rich game worlds that allow players to do a wide range of different activities, such as fighting, making friends, building, gathering, and so on. In this section, we will present a concept linked to games that we named "Game Path", which regards the sequence of players' choices inside the game content (i.e., how a player consumes the game content).

First of all, we assume game content as all available interactions between a player and the game environment. According to a player's advancement in the game story, the number of possible interactions may increase or decrease. For example, in a specific game, when a player achieves level 20, he/she can join PvP battles, when the player's intelligence attribute is over 100, his/her avatar can fly, or when the player changed his/her job from farmer to blacksmith, he/she cannot grow crops anymore but can forge weapons now. Therefore, the changes in the total amount of available interactions are specific for each game and for each player. Moreover, those changes can happen differently depending on the game story being linear or non-linear (ROLLINGS; ADAMS, 2003).

To better clarify the concepts of linear and non-linear stories, we will assume a game as a graph with a starting point and an ending point corresponding to the beginning and the end of its story. An example of a linear story is depicted in Figure 9. The red lines represent the main path, the path that leads to the ending point. The yellow lines represent optional paths, which do not lead to the ending point but can give to the players some additional rewards. Each node of the graph is a possible interaction (an action). Labels were added to identify uniquely each node where "Posx" regards the position "x" and "Opty" regards the option "y". In this representation, the "Opt0" always represents the main path (red line). Moreover, a player "walks" in this graph as he/she is playing (i.e., making choices), and at each position, he/she has a set of possible interactions (i.e., possible choices). For example, at Pos1, a player can acquire a quest, join a guild, or upgrade an item. The depth of both the main and the optional paths can vary depending on the considered game; the same happens to the number of options at each position. Moreover, some actions can be done more than once, while others are limited to only one accomplishment.

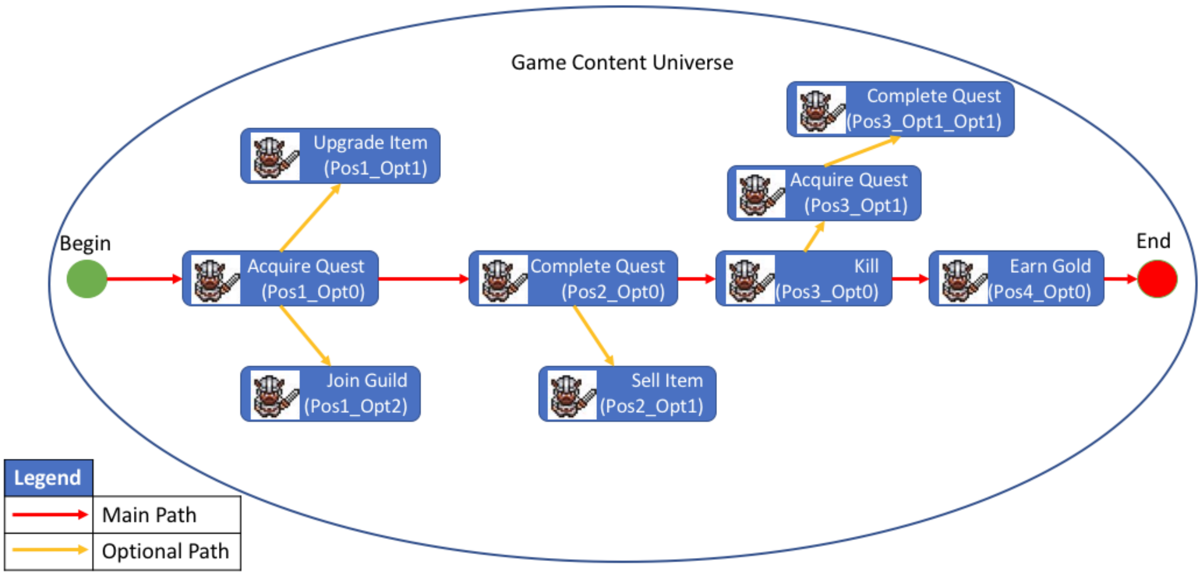


Figure 9 – Example of linear game story

The non-linear story representation is a little bit different, as we can see in Figure 10 (as this graph has more nodes than the linear story perspective and to better highlight the structure, labels for position and option were omitted on purpose as their concepts were already explained in Figure 9). In this situation, a player can be at more than one position at a time, and there is more than one main path. Despite the difference, the starting and ending points still exist. The non-linearity is considered a very desired feature ("Holy Grail") by game designers as it keeps players motivated for longer due to the wide range of possible interactions (ROLLINGS; ADAMS, 2003).

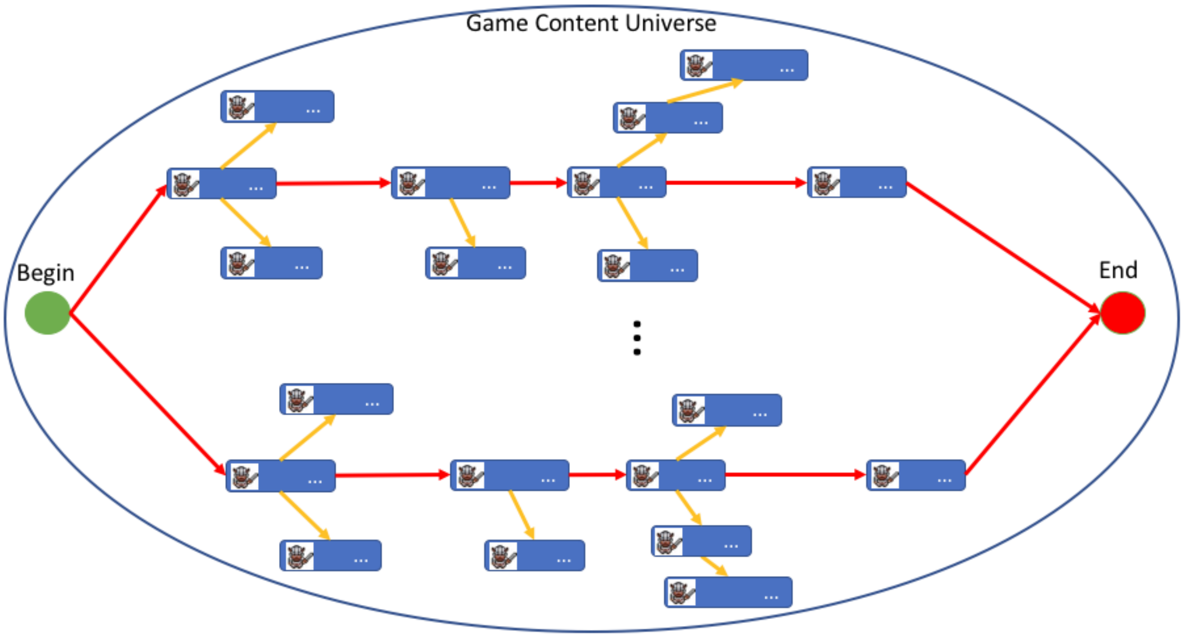


Figure 10 – Example of non-linear game story

Regardless of whether a game is linear or non-linear (with complex paths or not), the "Game Path" can be illustrated without the concepts of main path and optional path, as being a simple ordered list of actions. For example, assuming the simple game story of Figure 11 where a player did the following actions according to this order: Pos1_Opt1, Pos1_Opt0, and Pos2_Opt0. The same actions can be illustrated as the Game Path portrayed in Figure 12. This representation of Game Path will be explored in more detail in Chapter 5.

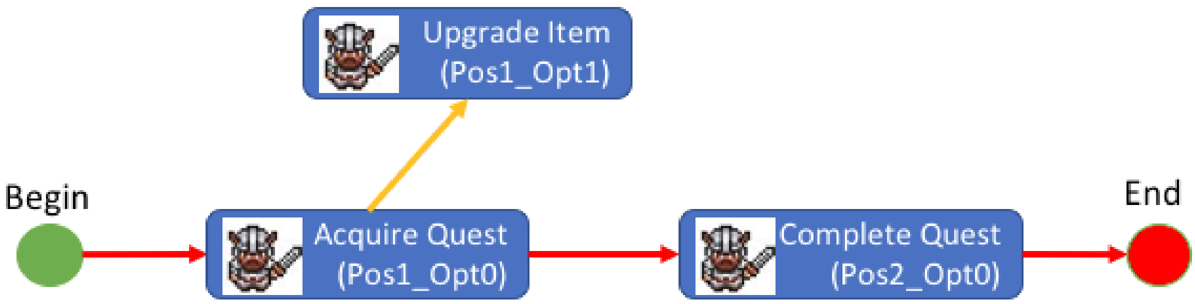


Figure 11 – A simple game story

Game producers may periodically release new game content, expanding the range of possible interactions to motivate players again. This kind of strategy postpones the end of the game story. In the case of the World of Warcraft (WoW), this strategy happened

Figure 12 – A simple Game Path

successfully for 16 years (MCALOON, 2018d) (with few back steps (KUMMER; NIEVOLA; PARAISO, 2017b)). If the WoW game producer did not provide new game contents, the game would probably "live" for much less time (CROSS, 2018).

In conclusion, we assume that the players' choices inside the game content are based on psychological aspects. Thereby, the more a game is non-linear, the bigger the range of options, allowing players to choose what they more like to do, avoiding what they do not like. These possibilities of choice foment that MMORPGs are an excellent option for a game genre to identify players' psychological profiles due to their wide range of possible interactions. In addition to it, other game genres may present additional challenges to identify such aspects since it might be offered fewer possibilities to the players compared to the RPG one, such as in the First-person Shooter (FPS) case, where basically, players can shot, reload, and run [6].

3.4 Risk Situations

After a game is released in the market, its usage lifecycle begins, and attached to it, the possible occurrences of risk situations. A risk situation is any situation that demands some actions to be done by the game producer, as it can impair the lifecycle longevity. The game producers' main concern is profit, so the longer a game "lives", the bigger the profit. Therefore, identifying and solving risk situations as soon as possible have great importance. Some examples of risk situations are:

- When players abandon the game after an initial try because their expectations were not met ("The Chasm" (MOORE, 1995) and frustration aspects (ZHU; LI; ZHAO, 2010)).
- When players start to have a lack of motivation due to the absence of new game content (new challenges). It happens due to the fact that when a game content is consumed, the game gradually loses its power of keeping players motivated (ZHU; LI; ZHAO, 2010; COOK, 2007).
- When the abandonment rate is greater than the new players rate (SPELLER_III, 2012).

[6] Nevertheless, featuring engineering can be applied to sequences of players' actions in FPS games to generate more abstract actions, like invading the enemy base, helping others, or scouting. More details about feature engineering can be found in Subsection 3.6.2

- When the new game content does not please the players (KUMMER; NIEVOLA; PARAISO, 2017b).
- When a game is entering the last stage of the usage lifecycle (Niche) (KUMMER; NIEVOLA; PARAISO, 2018b; KUMMER et al., 2016).
- When profitable players have a low expectancy of survival time (the expected amount of time that a player will continue to play, usually measured in days) (LEE et al., 2018; KIM et al., 2017; TAMASSIA et al., 2016; PERIÁÑEZ et al., 2016; RUNGE et al., 2014; KUMMER; NIEVOLA; PARAISO, 2018a).

The identification of risk situations is mainly made through analyzes of usage metrics, being MAU the most used (Equation 1.1). Next, examples of decisions made to mitigate or solve risk situations are presented:

- Generation of new game content (KUMMER; NIEVOLA; PARAISO, 2017a).
- Generation of a new game (upgrading the game mechanisms) (KUMMER; NIEVOLA; PARAISO, 2017a).
- The application of a recommendation system that suggests to players game content linked to their preferences (BERTENS et al., 2018).
- Contacting a player when he/she starts to show lack of motivation (SPELLER_III, 2012; MÜNTNER, 2017; KUMAR; SHAH, 2004).
- To add a clan (guild) system to increase the long-term retention (ANANKIN, 2018).
- To offer clear core-loops (repeatable challenges) with enough decision-making to improve mid-term retention (ANANKIN, 2018).
- Finishing the game lifecycle to avoid worse losses (KUMMER; NIEVOLA; PARAISO, 2017a).

In particular, we depict some additional comments about the decision to generate a new game. In this situation, the named "self-cannibalism" may occur, consisting of migration of players from the old game to the new one, forcing the ending of the old version (this process is illustrated in Figure 13). The good aspect is the assurance of active players in the new game, and the bad aspect is the contribution to end the old one.

The MAU metric can hide some important information because even though a player is playing, it does not mean that he/she is enjoying or motivated to play longer. The MAU metric can be considered a "raw" metric, as it considers only usage time. In the work of (KUMMER; NIEVOLA; PARAISO, 2017b), the authors proposed a metric called

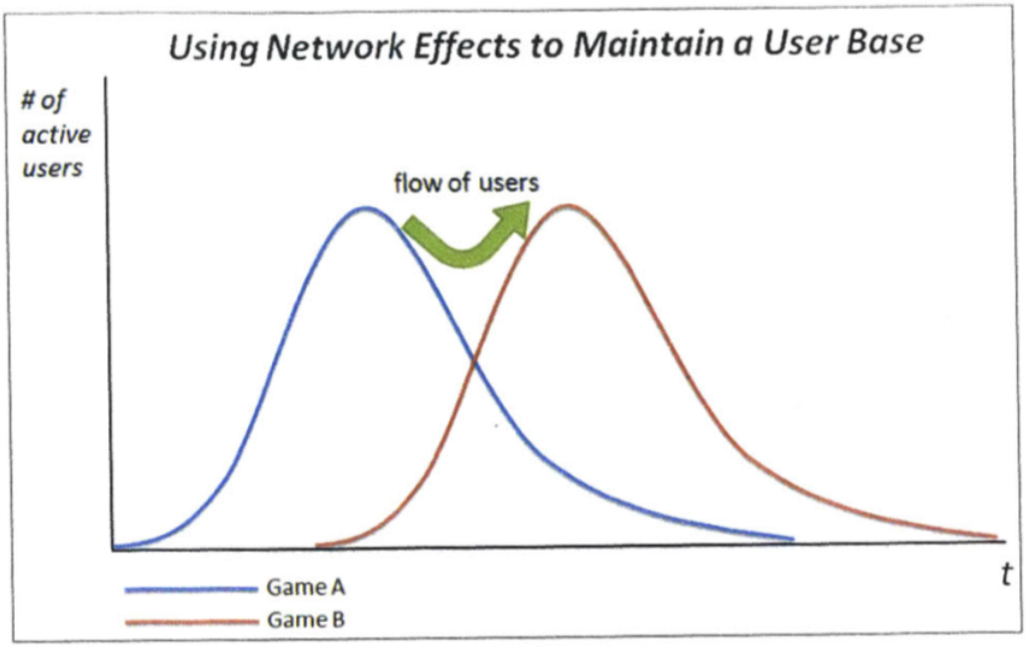


Figure 13 – Example of self-cannibalism, extracted from (SPELLER_III, 2012)

Commitment, which measures the motivation of the active players into three different degrees (low, average, and high) based on the amount of played time and obtained score. Moreover, this metric is obtained through a Machine Learning approach (Machine Learning is part of the KDD subject, which is explained in Section 3.6; more details about the Commitment computation can be seen in Section 3.7). Figure 14 depicts the MAU and the Commitment behavior according to game upgrades of World of Warcraft.

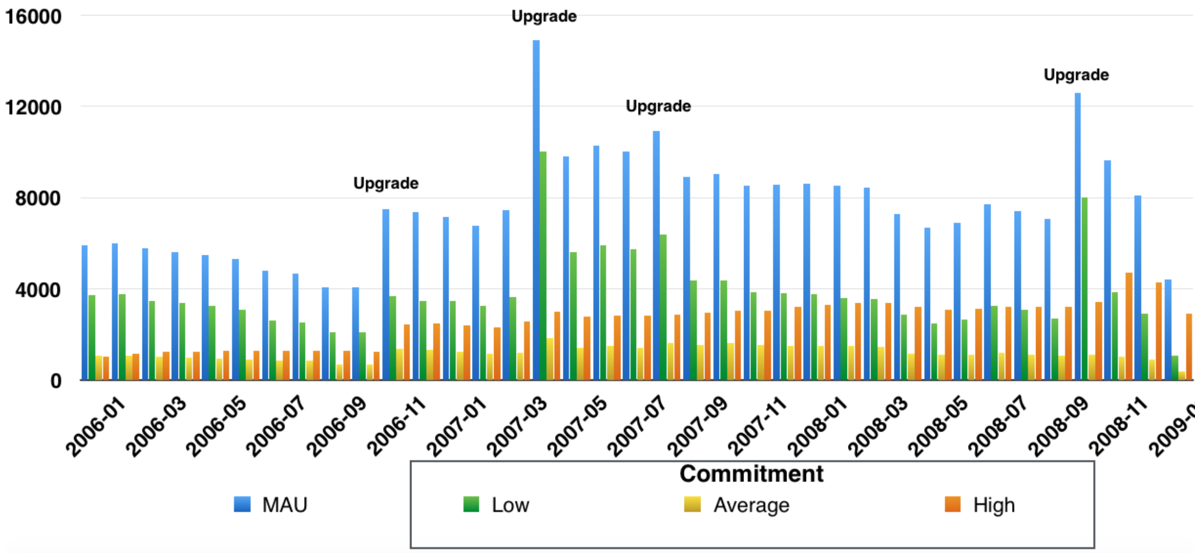


Figure 14 – MAU and Commitment comparison of WOWAH dataset, extracted from (KUMMER; NIEVOLA; PARAISO, 2017b)

As we can see in Figure 14, even though other months have similar MAU values, the Commitment arrangement may differ. It means that the behavior of the active players is changing over time. In addition to it, the four depicted upgrades show the growth of MAU. It indicates that the game upgrade created expectancies in players' minds, entailing in a wishing to explore the new content. The Commitment changes over time are better presented in Figure 15.

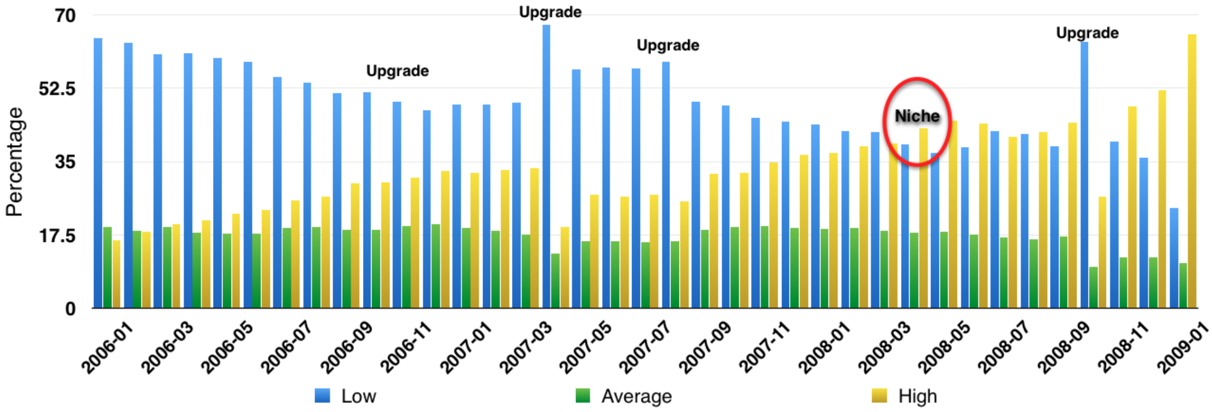


Figure 15 – Commitment percentage changes of WOWAH dataset, adapted from (KUMMER; NIEVOLA; PARAISO, 2017b)

According to the percentage representation of Figure 15, it is possible to notice a trend of reducing low committed players (new players) and increasing high committed ones. Moreover, when there is an upgrade, the distance between low and high percentages becomes bigger again, until the moment when the number of high committed players becomes greater than the number of low committed ones. When this situation is happening, the game is considered as being in the Niche stage (depicted in Figure 3 (COOK, 2007) and identified on usage data by (KUMMER; NIEVOLA; PARAISO, 2018b; KUMMER et al., 2016)).

A risk indicator was proposed by (KUMMER; NIEVOLA; PARAISO, 2017b) based on changes of players' commitment degree. As a player plays, he/she may change his/her commitment to the game, for example, from low commitment in the first month to average commitment in the second one. Thus, a player can increase or decrease his/her commitment over time. The authors computed those changes and proposed the RI (risk indicator) to evaluate the success or not of game upgrades. The indicator has a range between 0 and 1, where 1 means the series's best commitment growth. An interesting finding can be seen in Figures 16 and 17.

On the one hand, if we only look at the MAU metric (Figure 16) we may conclude that all upgrades were successful because the MAU grew up, but on the other hand, if we look at the RI value (Figure 17), we can see that in the last upgrade the players' commitment continued to drop. It means that the upgrade was not successful because even though there are more active players, they are less committed to the game and will leave soon, as can be checked in the last three months, where the MAU dropped continuously.

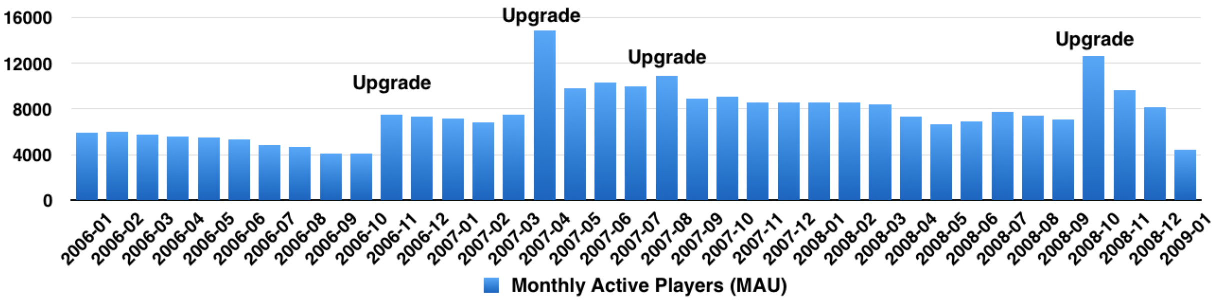


Figure 16 – MAU obtained from WOWAH dataset, extracted from (KUMMER; NIEVOLA; PARAISO, 2017b)

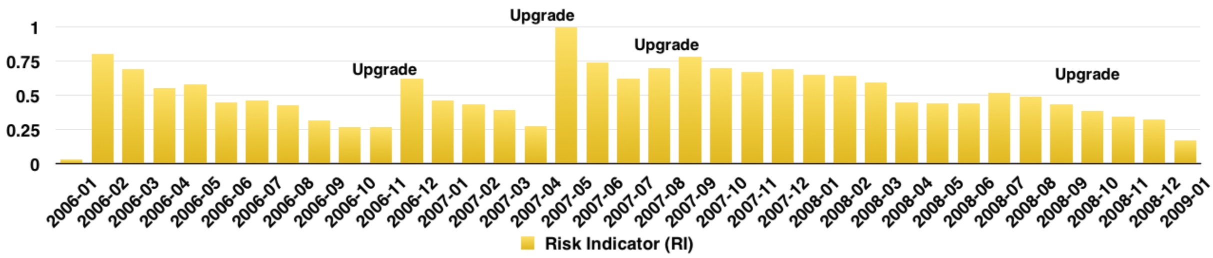


Figure 17 – RI obtained from WOWAH dataset, extracted from (KUMMER; NIEVOLA; PARAISO, 2017b)

If a game producer used only the MAU metric, the risk situation would be identified only three months after the last upgrade because the MAU value becomes lower than its value before the upgrade. However, if the game producer used the RI value, the risk situation could be identified in the upgrade month, giving precious time to improve the chances of solving the problem. Moreover, the best RI and MAU values regarded the same upgrade, which means that this upgrade was successful because it acquired new players and entertained the active ones.

A usage lifecycle is not made only of risk situations, good situations also happen. When a game has a good acceptance and profit, game producers may choose to improve the profit expanding the business to other platforms, such as consoles, PCs, smart-phones, tablets, among others (SHEFFIELD; ALEXANDER, 2008; SPELLER_III, 2012; GRAFT, 2009). Moreover, an example of a decision made over the entire lifecycle of a game is the “chase” for new players, which is usually done through advertisements (SHEFFIELD; ALEXANDER, 2008; SPELLER_III, 2012).

3.5 Game Refinement Theory

The Game Refinement Theory was initially proposed by (IIDA; TAKESHITA; YOSHIMURA, 2003) to measure a degree of interest (excitement) in board games. This theory avoids the empirical and subjective concept of interest, founding its concepts on the classical mechanic physics of Newton (NEWTON, 1687). In the same way that a roller-coaster can provide more or less fun according to settings of G-force, speed, and height, a board game may have a similar structure, as we have interests in both kinds of activities. On the one hand, the roller-coaster applies an acceleration into our body, which affects our sensation of pleasure. On the other hand, a board game also provides such sensation which can be the result of an acceleration, but in this case, applied to our minds. This is the essential idea behind the Game Refinement Theory (physics in mind).

Iida et al. (2003) modeled the acceleration in mind as the uncertainty of the game outcome. The assumption is that in exciting games, the game's final result must be unknown until the end of the game. In addition to it, games with this characteristic can be named as seesaw games, as an analogy to the changes on advantage from one player to another (uncertainty about the winner). Based on the seesaw concept (IIDA et al., 2004; IIDA, 2003), a model of game uncertainty was proposed.

From the player's point of view, the comprehension of the game result is a function of time (number of moves) t, as the result becomes more determined as time passes. Therefore, the amount of solved uncertainty can be represented by the function $x(t)$. This function represents the game information progress, which states how certain is the result of a game in a certain time. Let B and D be the average branching factor and the average depth of a game, respectively. In board games, the branching factor is the number of possible moves at a time t, while depth is the total number of moves until the end of a match. As the depth may vary according to the opted strategy (the selected branch), B and D's average values are used. If B and D are known for a match, the game information progress $x(t)$ will be given as a linear function of time t with $0 \leq t \leq D$ and $0 \leq x(t) \leq B$, as presented in Equation 3.1.

$$\mathrm{x(t)} = \frac{B}{D}t \tag{3.1}$$

However, the game information progress 3.1 is usually unknown during a match. Hence, it is assumed to be exponential due to its uncertainty until the very end of a game. Therefore, a more realistic model is given by Equation 3.2.

$$\mathrm{x(t)} = B(\frac{t}{D})^n \tag{3.2}$$

Where n is a constant given by an observer of the considered game. The acceleration of the game information progress can be obtained applying the second derivative of Equation 3.2. Solving it at $t = D$ (the end game period).

$$\mathrm{x(D)"} = \frac{Bn(n-1)}{D^n}D^{n-2} = \frac{B}{D^2}n(n-1) \tag{3.3}$$

It is assumed that when a match is happening, the acceleration of game information is happening somehow in our minds, being it enjoyable or not. The physics in mind is not yet fully understood, but according to Newton's laws, if there was an acceleration, there was a force acting, so when there is an acceleration in our minds, there is also a force acting. Therefore, it is expected that the larger the value $\frac{B}{D^2}$ (as n is a constant, the clause $n(n-1)$ can be omitted), the more exciting a game is, due to the outcome uncertainty. Thus, the Game Refinement Value (GRV) is assumed as described in Equation 3.4 (IIDA et al., 2004).

$$\text{GRV} = \frac{\sqrt{B}}{D} \tag{3.4}$$

It is expected that the bigger the GRV, the more entertaining a game will be due to its uncertainty until the very end of a match. In (IIDA; TAKESHITA; YOSHIMURA, 2003), the authors measured the GRV for traditional board games, details are presented in Table 16.

Table 16 – Measures of GRV for traditional board games, extracted from (IIDA; TAKESHITA; YOSHIMURA, 2003)

Board Game	B	D	$\frac{\sqrt{B}}{D}$
Chess	35	80	0.074
Xiangqi	38	85	0.073
Go	250	208	0.076
Shogi	80	115	0.078

Game Refinement Theory was also successfully applied to continuous movement games, score limit games, crane games, fighting games, RPGs, RTS games, and multiplayer online battle arena (MOBA) games, among others (as depicted in Table 17) (SUTIONO; PURWARIANTI; IIDA, 2014; XIONG et al., 2014; PANUMATE; XIONG; IIDA, 2015; PANUMATE et al., 2015; PANUMATE; IIDA, 2016a; XIONG; IIDA, 2014; XIONG; ZUO; IIDA, 2014; CHETPRAYOON, 2016; XIONG et al., 2015).

The key concept to apply the Game Refinement Theory to a game consists of identifying the game information progress model. For example, in the Soccer case, the average values regarding the number of shots and the number of successful shots are used; whereas in the Baseball case, the average values for hits and scores are used; while for the RPG case, the values of options available and turns can be used. As we can see, all of these values are applied in some way to the structure proposed in Equation 3.4.

A standard agreement of all applications regards the called "sophisticated zone" or "noble uncertainty" (YICONG et al., 2019), which is the GRV between 0.07 and 0.08.

Table 17 – A summary of GRVs for many different games

Games' Type	**Game**	**GRV**
Continuous Movement	Basketball	0.073
Continuous Movement	Soccer	0.073
Continuous Movement	Boxing - World boxe association (WBA)	0.085
Score Limit	Badminton - Old scoring system	0.121
Score Limit	Badminton - Current scoring system	0.086
Score Limit	Baseball - 2015	0.063
Score Limit	Softball - 2015	0.08
Crane Game	UFO Catcher - Japan	0.075
Crane Game	UFO Catcher - Thailand	0.057
Fighting Game	Super Street Fighters 4	0.0716
Fighting Game	The King of Fighters 98	0.1041
Fighting Game	The King of Fighters 13	0.1149
RPG	Pokemon 6th generation Catching (average value)	0.065
RPG	Pokemon - Battle (human data)	0.058
RPG	Pokemon - Battle (simulated data)	0.061
RPG	Pokemon Red 1st generation Gameplay	0.072
RTS	StarCraft II Average value between all races	0.074
MOBA	Dota 2 - version 6.8	0.078
MOBA	Heroes of the Storm Average value between all maps	0.092

Games in this zone tend to be more enjoyable than games outside of it. Moreover, values lower than 0.07 means that a game is more competitive (more based on players' skill), while values above 0.08 means that a game is more entertaining (more based on chance; where a less skilled player can win); therefore, the sophisticated zone is a balance between players' skill and chance. Another aspect regards the game length, where games too long can be seen as boring (lower GRV entailed by a bigger D value) while games too short as unfair (higher GRV entailed by a smaller D; firmly attached to chance). An interesting finding is that the historical evolution of games shows convergence to the sophisticated zone (YICONG et al., 2019).

The GRV is usually obtained to a game as an overall perspective (average values), but nothing prevents to obtain it for individual matches (case-by-case) (PANUMATE; IIDA, 2016b). Moreover, the same game may have many GRVs, as we can see in the Pokemon case in Table 17 (values for catching, fighting, and general gameplay).

3.6 Knowledge Discovery in Databases (KDD)

The Knowledge Discovery in Databases is a field of study that embraces Artificial Intelligence, Statistics, Machine Learning, Pattern Recognition, and Database (DB). The KDD's primary motivation is the identification of useful information that was not previously available ("hidden"), which may be interesting for some application (e.g., the identification of profitable costumers and their tendencies, the identification of malicious users, and the occurrence or not of a disease). The discovered knowledge is usually used to avoid empirical considerations, helping decision-makers base their choices systematically.

In the course of time, studies and algorithms were developed, applied, and evaluated. The process to do the KDD is usually divided into the following steps: data selection, preprocessing, transformation, data mining, evaluation, and interpretation of results (TAN; STEINBACH; KUMAR, 2005). Next, the steps considered in this research are explained.

3.6.1 Data Selection

The data selection is the first step of KDD, but there is no pattern for its appliance; each case is considered at a time. Let us assume as an example an MMORPG that has on its usage data values of avatars' level, name, money, and current place. Assuming that a game producer wants to identify the sequence of places that players visit the most, and then induce a model capable of receiving as an entry the avatar's current place and giving as an output the most likely next place. The information regarding avatars' name, money, and level may be irrelevant, so only the current place data would be selected.

This step is crucial to the KDD process, as it defines the data universe in which the discovery will happen. If relevant information does not exist in the selected data, the final result may not be satisfactory or does not present useful information. Given it, the understanding of all elements present on the available data must be done, allowing in that way the elaboration of justifications to use each one or not. In our example, if the avatars' current place was not selected, the final result could be unsatisfactory.

In the game context, an instance of the usage data is the term used to identify the set of characteristics (attributes) of an avatar at a given time. It is assumed that instances are not duplicated in the database. Each instance's unique identification can be a numerical code or the combination of the avatar's name (or ID) with the respective time.

3.6.2 Preprocessing and Transformation

The preprocessing and transformation steps are similar. In some situations, the original data format is not appropriate for the data mining algorithms (the next step); therefore it is necessary to preprocess it. The data collection procedure may occur in different ways, such as by sensors, manual actions, or queries in DBs, but independent of it, failures may happen. Missing values (e.g., an avatar without a name) or incompatible values (e.g., a timestamp in the avatar's level attribute) can exist. Thus to solve any possible problem, a preprocessing must be performed to fix all inconsistencies.

Another situation regards the transformation of the original data. Let us assume that a game producer wants to check if new players enter a determined map of a game. However, there is no attribute specifying if a player is or not a new player. In this case, the avatar's level may be used as a reference. Assuming the range of levels between 1 and 60, one may divide it into three zones, where the first zone references the range between 1 and 20 (the new players' zone), the second zone between 21 and 40 (average players' zone), and the third and last zone between 41 and 60 (expert players' zone). After applying this transformation to the original usage data (that can be already preprocessed), the supposed concern can be managed. Despite the proposed example, there are other transformation techniques in the literature that varies in complexity (TAN; STEINBACH; KUMAR, 2005; SKANSI, 2018).

In particular, we highlight a transformation technique named "one-hot encoding" (SKANSI, 2018), which is adopted in some of this thesis's experiments. This technique is useful when the adopted data mining algorithm does not accept nominal features but only numerical ones. By applying this technique, nominal features can be transformed into numerical ones following a boolean idea. Assuming the next feature, "avatar's hair color", represented in Table 18, the execution of the one-hot encoding generates as outputs the values presented in Table 19. As we can see, the same information is present but portrayed mathematically by two possible values, 0 for False and 1 for True. Note that if there are 40 possible hair colors, the one-hot encoding will generate 40 features.

Table 18 – Example of nominal features to be processed through the one-hot encoding technique

Player ID	**Avatar's Hair Color**
1	Blue
2	Red
3	Orange

Table 19 – Example of nominal features transformation through the one-hot encoding technique

Player ID	**Avatar's Hair Color - Blue**	**Avatar's Hair Color - Red**	**Avatar's Hair Color - Orange**
1	1	0	0
2	0	1	0
3	0	0	1

An additional aspect of the preprocessing and the transformation is the called "Feature Engineering". In some cases, the data format is too "raw" to give some useful information to the data mining algorithms; therefore, more accurate information is needed. An example of that is the work of (KUMMER; NIEVOLA; PARAISO, 2018a). In this work, the authors counted the number of occurrences of each action of each player in an MMORPG and then computed a called "Tendency metric", which represents the growth or decay tendency of each action (note that this approach gives more information than just a boolean value stating if a player did or not an action). Their approach was compared with others in a Data Mining Competition (LEE et al., 2018) and obtained first place on the two considered aspects (the prediction of players' churn and survival time). It is interesting to notice that their approach did not apply advanced Data Mining techniques (such as Deep Learning[7]), but instead of it, they focused on feature engineering and obtained better results (statistically relevant) than more advanced Data Mining approaches. This fact highlights the importance of giving (as much as possible) useful information to the Data Mining algorithms. Note that many different approaches can be applied to the same dataset (entailing in different new information), but independent of that, the focus is on identifying useful information. It is possible to say that the feature engineering process is more empirical than systematic because there is no concrete evidence that an approach will get better results. Therefore, the assessment of a new feature as good or not is done through experiments, usually comparing two datasets' performance, one with the new feature and another without it.

3.6.3 Data Mining and Evaluation

This step regards the application of algorithms in the preprocessed and transformed data (if they were needed). There are different categories of algorithms, each with a specific purpose (linked to the called mining activity (TAN; STEINBACH; KUMAR, 2005)). Classifiers label instances of future data based on classes learned on historical data. Regressors are similar to classifiers, but instead of predicting a label, they predict a numerical value. Cluster approaches aim at identifying groups with similar characteristics (assuming a "distance" measure between instances). In addition to it, the appliance of such algorithms is restricted to the characteristics of the dataset. For the supervised approaches (e.g., classifiers and regressors), a label (class) or numerical value is needed apriori, what does not happen to the unsupervised ones (e.g., clustering).

Regardless of the chosen algorithm, there are similar characteristics; for example, the necessity of historical data. In the case of classifiers and regressors, they predict future behavior based on the historical one, while cluster approaches use it to identify groups and predict future instances. Creating a model based on historical data is named Induction

[7] This subject is approached in Subsection 3.6.3.3.1.

(the model is induced from the data). The use of a model is called Prediction (the model predicts results for future data). As an additional remark, in cases where a supervised approach is desirable, but no labels (classes) are available, the identified groups of a clustering approach can be assumed as classes.

Every induced model represents a hypothesis about how to understand the data. For the same data, many different models can exist (i.e., many different points of view). Besides this fact, some metrics exist to evaluate the performance of a model. These metrics can change according to the chosen algorithm category. Next, the approaches applied in this work are presented.

3.6.3.1 Classifiers

Classifiers aim at labeling instances with a class. A class is a nominal value that already exists in the historical data. For example, a class could be the commitment degree of a player to a game (low, average, or high) (KUMMER; NIEVOLA; PARAISO, 2017b).

There are many strategies (algorithms) to predict classes to instances, being possible to divide them into two great categories, the "black-box" and the "transparent-box" ones (WITTEN et al., 2016). On the one hand, transparent approaches are more comfortable for humans to understand because they illustrate how the result was obtained. On the other hand, black-box approaches do not have such property. Depending on the considered problem, transparent box algorithms can be used to show to an interested person (usually the one who makes decisions) how the result was generated, as it is common to have new information that counters the common sense.

Decision trees and rules are examples of transparent-box approaches (TAN; STEINBACH; KUMAR, 2005). In the decision trees case, each leaf node is a class, and each non-leaf node is a point of decision that uses an attribute (characteristic or feature) value to decide which node (branch) to follow (only one path is chosen). Figure 18 shows an example of a decision tree that predicts players' commitment based on avatars' level and amount of played time (in days, according to a given month). Decision trees can be translated to rules and vice-versa, using the Figure 18 as an example, the rule to define low committed players is "if an avatar has less than (or equal to) 25 days played and its level is lower than (or equal to) 30, then the player has a low commitment to the considered game".

The higher (or nearest to the root) a non-leaf node is, the more relevant its attribute is. Moreover, if an attribute is not relevant, it may not be used by the algorithm. In Figure 18 example, the amount of played time is more relevant to distinguish different classes than the level attribute. An example of a measure that is used to allocate an attribute into a node is the "information gain" (QUINLAN, 2014).

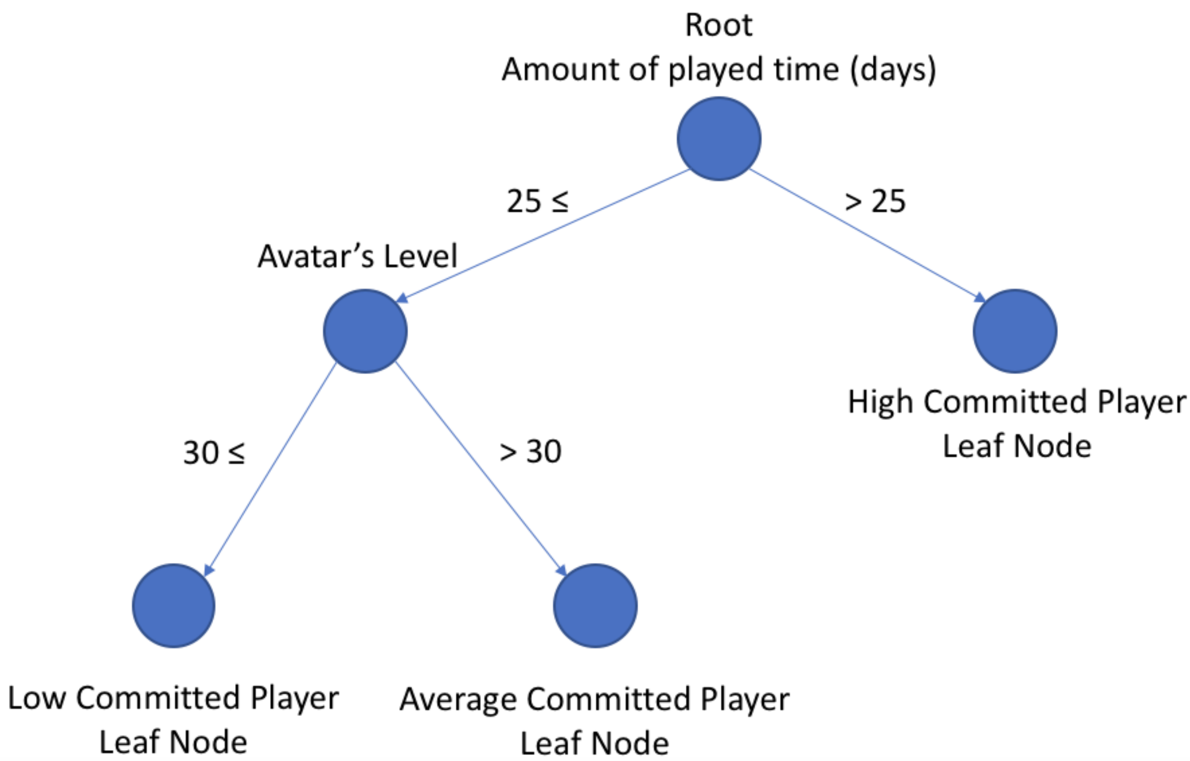


Figure 18 – Example of a Decision Tree structure

Neural networks and SVM (Support Vector Machine) are examples of black-box approaches (TAN; STEINBACH; KUMAR, 2005). Neural network models are induced through many iterations in the same DB. Where at each iteration, some neurons' entries may have their weights changed depending on whether the final result was predicted correctly or not. A neuron is the most basic structure of a neural network and is based on the biological neuron. Therefore, it has one or more dendrites (entries), a core (processing), and one axon (that connects this neuron to others; result propagation). In the Data Mining approach, each dendrite refers to an attribute and has an associated weight. Every time a neuron misses, the weights are adjusted; this explains the necessity to do many iterations until a feasible performance is obtained. A neural network then is a set of neurons connected to each other obeying a predefined structure (there are many possible configurations (SKANSI, 2018)). SVMs use the addition of a hyperplane to divide instances into groups (classes). In both approaches, the algorithms' internal operation is mathematically represented, turning it less easy to be understood by a human.

Regardless of the algorithm chosen for the classification problem, a model is always created based on historical data. This creation has two main phases, the training (induction) and the testing (prediction). A classifier objective is not to predict with 100% of accuracy the historical data, but to use the historical data to induce models capable of predicting the future data with the best accuracy as possible. An always-present challenge in the Data Mining approaches is that the data behavior changes over time; therefore, what was learned in the past may be wrong in the future. Moreover, a perilous situation is when a model is firmly attached to the historical data patterns (overfitting), as simple changes in future behavior can drastically impair the classification performance. At least, a classifier must have an accuracy greater than 50%; otherwise, the use of a random algorithm could have a better performance. To avoid the problem of overfitting, the historical data can be divided into two groups, one for training and another for testing. This division is usually done in two ways: a simple percentage division (e.g., the holdout approach, where usually 60% of data is designated for training and 40% for testing, however, other percentage values can be adopted, such as 80% for training and 20% for testing, following the Pareto's principle [8] (MIDDAUGH, 2015)) or cross-validation. The cross-validation approach divides the historical data into n parts (also known as folds), where $n - 1$ parts are used for training and one part for testing. It has a determined number of iterations based on n, where the iterations stop when each part was used as a test once. After this process, the final result is assumed as the average performance of all iterations.

A final classifier evaluation can be done through the use of metrics. Examples of such metrics are Accuracy, Recall, Precision, and F-measure (or F1-score). To better clarify these metrics' concepts, let us assume a simple example of 100 instances containing only two classes, a positive (50 instances) and a negative (50 instances). The final result can be represented by the number of true positives, true negatives, false positives, and false negatives cases. These cases are usually illustrated by a "Confusion Matrix", Table 20 shows the Confusion Matrix to our example.

Table 20 – Confusion Matrix example

		Actual class from historical data	
		Positive Class	Negative Class
Predicted Class	Positive Class	45 (True Positive)	15 (False Positive)
	Negative Class	5 (False Negative)	35 (True Negative)

Looking from a top-down perspective, the 50 instances of each class can be identified, while looking from the left-right perspective, the amount of prediction for each class is presented. In an optimal prediction case, only the diagonal value (from top-down and left-right) should have values greater than zero. Considering the proposed example, it is possible to see that the classifier did not predict the instances correctly for both classes, as five instances of the positive class were predicted as being of the negative class, and 15 instances of the negative class were predicted as being of the positive class. The correctly predicted instances represent the true positive (TP) cases for the positive class, the false positive (FP) cases are the ones where instances are predicted as the positive class when they are actually from the negative class, false negative (FN) cases are the ones where a positive instance is wrongly predicted as being of the negative class, and the true negative

[8] This principle states that 80% of the effects usually are originated from 20% of the possible causes.

(TN) cases are represented by the correct prediction of the negative class. Based on those concepts of TP, FP, FN, and TN, the Accuracy, Recall, Precision, and F-measure can be obtained. The Accuracy metric represents the percentage of correctly predicted instances (Equation 3.5), while the Recall metric expresses the correct identification of instances from the actual class (Equation 3.6), the Precision metric measures the correct identification of instances from the predicted class (Equation 3.7), and the F-measure is the harmonic mean between Precision and Recall (Equation 3.8). The metrics' values for our example are presented in Table 21.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.5}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.6}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.7}$$

$$\text{F-measure} = \frac{2TP}{2TP + FN + FP} \tag{3.8}$$

Table 21 – An example of evaluation metrics of classifiers

	Class		
	Positive	**Negative**	**Average**
Accuracy	80%		
Recall	0.9	0.7	0.8
Precision	0.75	0.875	0.81
F-measure	0.81	0.77	0.79

Note that all metrics can be computed for each class except by the Accuracy metric, just assuming the negative class as being the positive one. The Accuracy range is from 0% to 100%, where the nearer 100%, the better. The range for Recall, Precision, and F-measure are between 0 and 1, where values near 1 mean good performance. In addition to it, if there are more than two classes, the negative class can be assumed as a combination of all classes disregarding the positive one.

It is essential to highlight that, assuming a two-class problem (A and B), class A's wrong prediction may have a different impact compared to a wrong prediction of class B (these errors are different). Let us assume a situation where class A is the existence of a disease, and B is its non-existence. The impact of stating that a person has a disease while it is not true is different from stating that this person is not sick while indeed he/she is. Class A's wrong prediction may lead a person to do some unnecessary treatments, entailing in extra costs, while the wrong prediction of class B may lead a person to death.

Therefore, the penalty of erroneous identification of a class occurrence varies depending on the case. Linked to it, another critical aspect of evaluating a classifier is the relation between Accuracy and F-measure. Let us assume that the disease identification model has an accuracy of 99%, which can be considered very good, but the F-measure for class B (severe disease) is near zero. It means two things, (1) the number of class A occurrences is greater than the occurrences of class B, and (2) class B has a poor identification. In conclusion, even though a classifier has a good Accuracy, all F-measures of its classes must be checked together with the impact of miss-predicting each one because a serious problem may be happening behind the "good performance". A possible solution to this problem is the adoption of different weights for each class (applicable to neural networks), where the less present class can have a greater weight than the most present class, forcing the network to give more importance to learning the less present class nuances.

Another remark about classifiers regards the use of many classifiers to deal with the same problem. The idea is to take advantage of several different hypotheses (points of view) to obtain better performances; this concept is called "Ensemble of Classifiers" (TAN; STEINBACH; KUMAR, 2005). To predict a class for an instance, an Ensemble runs each classifier considering this instance, collects their predictions, and then applies a policy to decide the final predicted class. One example of a policy is the majority vote, where the most "voted" class is the chosen one. The evaluation process for Ensembles is the same as for individual classifiers. The ensemble creation process is adjustable, being possible to initialize it with classifiers previously induced, induce new ones based on a dataset, or combine both strategies. Moreover, inside an Ensemble, there is no restriction to have different classifiers (e.g., decision trees, rules, neural networks, or SVMs).

3.6.3.2 Regressors

Regressors are very similar to classifiers; the main difference regards their final result, which is numerical, different from the nominal one returned by classifiers. This difference entails another approach to evaluate the induced model, as there are no classes. Therefore, in this case, the final evaluation is given by a correlation value, which regards the tendency of the generated result accompanies the expected result.

Decision trees for classifiers and regressors work similarly, the leaf nodes that contain the classes for classifiers contain the numerical values for regressors, and in some cases, they may contain equations. The division of the historical data into training and testing follows the same rules as for classifiers.

3.6.3.3 Classifiers and Regressors associated with Time Series

An interesting nature of data is that the analysis of temporal changes in the values of its features (i.e., analysis of Time Series) can lead to better prediction. This kind of analysis can be achieved in two ways: (1) by encompassing the temporal changes through a feature engineering process (such as shown in Subsection 3.6.2 regarding the Tendency metric), or (2) by adopting a classifier or regressor that can receive as input Time Series.

Even though there are different approaches to work with Time Series, this thesis focuses on one that encompasses the capacity of considering both short and long-term aspects of features changes, the Long Short-term Memory (LSTM) network[9].

3.6.3.3.1 Long short-term Memory Networks

An LSTM network is considered a Deep Learning approach due to its internal architecture that contains recurrent connections between its improved neurons (i.e., the called LSTM cells or units). Such architecture encompasses the ability to consider past information or disregard it when it is not useful.

The conception of LSTMs was based on the continuous evolution of neural networks, that evolved since the idea of perceptron (a single neuron), passing through the concept of traditional feed-forward networks (layers of neurons), until the conception of recurrent neural networks (the addition of recurrent connections between layers or between neurons in the same layer). This evolution presented limits that were not entirely overcome so far, such as the problem of Gradient Vanishing. To understand this problem, first, it is needed to understand how a neural network learns. A neural network learns through the update of its weights, which is performed by the Backpropagation process. When the network misses, a value of error is generated. This value is used to update the weights from the network final layer until its first one according to derivatives founded on the called "Chain rule". The Gradient Vanishing problem regards the fact that the farther a layer is from the final layer, the lesser is its weights update. It means that, depending on the number of adopted layers, some layers can receive an update value so small that the network cannot learn. For an interested reader, we suggest this book (SKANSI, 2018) as an introduction to Deep Learning, which contains all the details and challenges since the conception of perceptrons until the aforementioned LSTM network. Next, the aspects of LSTMs used in this work are presented.

LSTMs are configurable through the called hyper-parameters, which defines how the LSTMs learning process will happen. In this thesis, the following hyper-parameters are considered:

- Learning rate: defines the pace that the network learns. Default values are usually 0.1, 0.01, or 0.001.

[9] LSTM networks can be applied to both classification and regression problems.

- Momentum: regards the amount of the last update to be considered in the current update.
- Batch size: defines the frequency in which the weights are updated. A batch size of 10 means that the weights are updated every ten instances.
- Validation dataset: a percentage of the training dataset can be reserved as a validation dataset. This concept is similar to the division of training and testing presented earlier in Subsection 3.6.3.1; the difference is that the validation set can be used to define when the training should stop. For instance, even when the training dataset error is decreasing, it is possible that the error in the validation set is increasing, indicating the beginning of overfitting and the right moment to stop training.
- Class weight: as previously mentioned in Subsection 3.6.3.1, it is possible to set the relevance of each class, such as performed by this hyper-parameter.
- Epochs: is the number of times that all training instances will be used in training.
- Activation function: is the internal function of a neuron that generates its output. Traditional functions return values between 0 and 1 (sigmoid), -1 and 1 (hyperbolic tangent), or the max value between an output x or 0 (rectified linear).
- Loss function: the loss function, or error function, regards the adopted equation that will generate the error value in case of a network miss. For problems with two classes, which is the case of the further experiments, the "Binary cross entropy" is a valid option, as it computes the error considering the perspectives of both classes, such as shown in Equation 3.9.

$$\text{BinaryCrossEntropy} = -\frac{1}{N}\sum_{i=1}^{N} y_i * log(p(y_i)) + (1-y_i) * log(1-p(y_i)) \tag{3.9}$$

Where N is the number of instances, y_i the actual label of instance $_i$, and $p(y_i)$ is the probability of $_i$ being predicted as y. In sum, this equation considers the log probability of an instance being from one class or the other.

The combination of a lower learning rate, a 0 momentum, and the use of a validation dataset is suggested to prevent overfitting. It is essential to highlight that there is no perfect procedure to follow that generates as a final result the ideal hyper-parameters to a given problem. Therefore, an acceptable configuration of hyper-parameters must be chased through experiments. Nevertheless, there are some traditional rules called "rules of thumbs"[10] that guides the configuration of hyper-parameters, however, without a systematic procedure.

3.6.3.4 Clustering Approaches

The concept of clustering differs from the ones for classifiers, and regressors (TAN; STEINBACH; KUMAR, 2005). In this perspective, the historical data do not have any labels or numerical values that allow identifying an instance's group (or cluster) apriori. Therefore, the clustering task's objective is to identify similar behaviors based on the instances' attributes values, using them as criteria to create groups.

There are many strategies to identify groups. A common approach is the one portrayed in the k-means algorithm (ARTHUR; VASSILVITSKII, 2007). In this approach, the called K value represents the number of groups to be identified (which is settable). The algorithm starts plotting into an n-dimensional hyperspace (where n regards the number of attributes) all instances based on their attributes' values. Then, K centroids are plotted randomly in the hyperspace, and after that, the iterations start. At each iteration, each instance is associated with the nearest centroid; afterward, each centroid's position is adjusted as the mean position of all its instances. This process continues until a limit of iterations is reached, or there are insufficient changes between the last iteration and its predecessor (obeying a predefined threshold).

The definition of distance between an instance and a centroid can be done through different measures, for example, applying the Euclidean distance (i.e., measuring the distance between two points using a straight line; Equation 3.10).

$$\text{EuclideanDistance} = \sum_{i=1}^{n} \sqrt{(p_i - q_i)^2} \tag{3.10}$$

Where n is the number of attributes, p is the position of an instance, q is the position of a centroid, p_i is the value of the $_i$ attribute of the instance p, and q_i is the value of the $_i$ attribute of the centroid q. The evaluation of a clustering induction can be done by summing up each instance's distance to its centroid (the cluster "error") and then summing up all the clusters' errors. This approach is called Sum-of-Squared-Error (WALDE, 2006).

The division of the historical data into training and testing follows the holdout concept, where after the final centroids' positions are defined, the testing instances are plotted, and then an evaluation metric is applied. It is essential to highlight that the cross-validation concept does not work well for clusters because there is no correct answer for an instance to check (the essence of unsupervised approach, i.e., no labels apriori).

[10] One example of such rules can be found at: https://towardsdatascience.com/17-rules-of-thumb-for-building-a-neural-network-93356f9930af .

In the k-means algorithm, the K value can be changed until an acceptable value of an evaluation metric is obtained (what is acceptable is usually given by a specialist). Another approach is giving as an entry historical data with the apriori known group (label) of each instance, and then validating if the identified group for each instance is the same as the actual one. Note that the Accuracy metric can be computed in this case.

The holdout approach is used to avoid overfitting (as described in subsection 3.6.3.1), but there is a particular case when 100% of the instances are used in the training set. It happens when one wants only to identify the current behavior and not predict the future one. Examples of such cases are the ones where a classification problem is demanded, but there are no classes available, then the clustering task is used as a "preprocessing" to identify the classes. Examples of this kind of application can be found in (KUMMER; NIEVOLA; PARAISO, 2017b; KUMMER; NIEVOLA; PARAISO, 2018b; KUMMER et al., 2016).

3.6.4 Interpretation of Results

KDD aims at identifying hidden and useful information to an interested person. So, after having the new information in hands, its assessment starts, usually with the interested person and a specialist. If the new results are understood as credible, the decision-making may begin (examples are described in subsection 3.4), and if it is not useful, it is discarded, and then another approach starts. It is essential to highlight that the conclusions about the new information can be unexpected (highlighting the occurrence of a good or bad situation), possibly demanding urgent actions.

3.7 Commitment Metric

The Commitment metric focus on identifying inside the active players, their degree of engagement to a game, being it obtained through a Machine Learning approach that was improved over time (KUMMER et al., 2016; KUMMER; NIEVOLA; PARAISO, 2017b; KUMMER; NIEVOLA; PARAISO, 2018b). The method assumptions are the following:

- The candidate game must be focused on entertainment (voluntary usage).
- The usage data must contain:
  - The obtained score
  - Timestamp
  - Players' ID

- The score must highlight the improvement of the players' abilities (e.g., in a soccer game, a player can win for 1x0, and after some time, win for 10x0 in the same conditions).

The method is based on usage data that are divided by time-spans, which can be daily, weekly, or monthly. Its steps are presented in Figure 19. For each time-span, unsupervised and supervised approaches are applied.

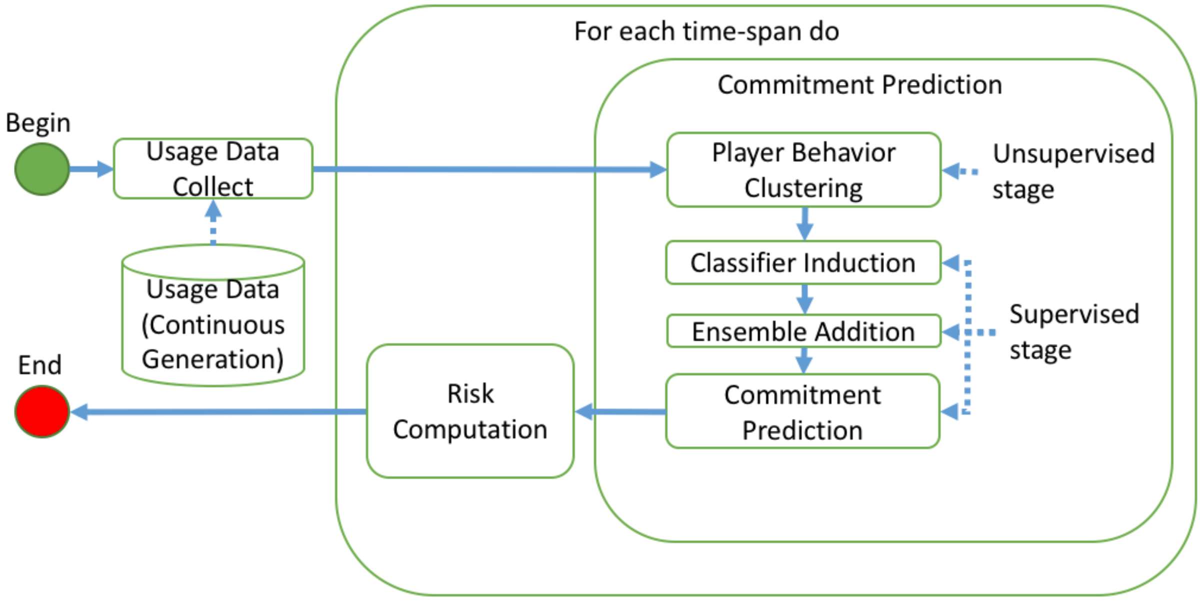


Figure 19 – Commitment identification method, extracted from (KUMMER; NIEVOLA; PARAISO, 2017b)

The unsupervised step ("Player Behavior Clustering") identifies, inside the usage data, three degrees of commitment (low, average, and high; for more details about the definition of three groups, please see (KUMMER; NIEVOLA; PARAISO, 2017b)) based on the minimum obtained score, the maximum obtained score, its range (*maximum* − *minimum*), and the total number of days played[11]. The clustering result is then used to induce classifiers in the "Classifier Induction" step, where each induced classifier is added to an Ensemble (in the "Ensemble Addition" step), which applies the majority vote policy to identify for each player his/her commitment degree to a game in a given time-span (the "Commitment Prediction" step). In an improved version of the method, only classifiers with different behavior are accepted in the Ensemble (KUMMER; NIEVOLA; PARAISO, 2018b). The final step ("Risk Computation") regards the identifications of the RI (Risk Indicator) value and the Niche stage (as depicted in Section 3.4). The Commitment metric for each time-span is obtained through a sum of all players on each commitment degree

[11] For time-spans with a monthly or weekly perspective, the number of days or hours can be used, while for the daily perspective, only the number of hours is functional.

(Equation 3.11, 3.12, and 3.13; examples of Commitment values are illustrated in Figures 14 and 15).

$$\text{Low} = \sum_{i=1}^{n} Plow_i \tag{3.11}$$

$$\text{Average} = \sum_{i=1}^{n} Pavg_i \tag{3.12}$$

$$\text{High} = \sum_{i=1}^{n} Phigh_i \tag{3.13}$$

Where n is the total amount of players, $Plow_i$, $Pavg_i$, or $Phigh_i$ has value 1 when the considered player $_i$ has low, average, or high on commitment respectively according to a given time-span, a value of 0 is assumed otherwise. It is essential to highlight that a player has only one commitment degree to a game in a given time-span, and also, a player can change his/her commitment degree to a game from one time-span to another one.

3.8 Introduction to Concept Lattices

Concept Lattices are knowledge representations that allow for improved analysis of linkages between concepts as it is founded on both psychological constructs (SOWA, 1983) and a mathematical basis (CARPINETO; ROMANO, 2004; GOLDREI, 2017). This thesis addresses and describes only some parts of the Concept Lattices, which are used for additional analysis of the Unification Explorer Framework presented in Chapter 4. For those with additional interest, we recommend reading this book (CARPINETO; ROMANO, 2004) for a deeper understanding of other Concept Lattices characteristics. The concepts applied in the current work are as follows:

- **Concept Lattice**: is a graph that follows the idea of context, where each node is a concept that can contain attributes and/or objects.
- **Context**: contains three elements (G, M, I) where G refers to objects, M attributes, and I the relations between G and M (i.e., the called incidence relations).
- **Object**: each object contains a set of attributes.
- **Attribute**: is a characteristic that an object can have or not. There is no limit to the number of attributes associated with an object.
- **Concept**: is a pair (A, B) of the context (G, M, I) where A and B refer to the extent and intent, respectively. According to Aristotle's rationales, essentially, a concept has the number of necessary attributes to describe itself (intent) attached to the

objects that it represents (extent). Objects that share the same set of attributes will be placed in the same concept.

- **Intent**: represents all of the attributes linked to a given concept.
- **Extent**: represents the objects linked to a concept.
- **Height**: the height of a concept is understood as the number of nodes from the maximal (the top) until the concept, being the bottom of the lattice called minimal.
- **Order**: attached to the height idea, there is the idea of order (from the top to the bottom of the lattice), where each concept that links to an upper concept inherits that concept's intents, while those linking to a lower concept inherit all of its extents.
- **Maximal**: a Concept Lattice has only one concept at its top, called maximal. Its height is assumed as zero. From the M perspective, if a Concept Lattice has a common attribute over all of its objects, this attribute will be placed on the maximal. By contrast, if there is no common attribute, the maximal will be empty. From the G perspective, if the Concept Lattice has an object without any attributes or containing only the attributes presented on the maximal, it will be placed on the maximal.
- **Minimal**: is the opposite of the maximal concept, being placed at the bottom of the lattice (there is only one minimal). From the M perspective, if a Concept Lattice has an attribute that is not linked to any of its objects, this attribute will be placed on the minimal. By contrast, if there is no unlinked attribute, the minimal will be empty. From the G perspective, if the Concept Lattice has an object containing all the attributes of the context, it will be placed on the minimal; otherwise the minimal will not have any objects.

In Figure 20, a simple Concept Lattice is presented using the transport context (containing three attributes and four objects), where a blue semicircle indicates the presence of one or more attributes in the concept (written with a gray background) and a black semicircle indicates the presence of one or more objects (written with a white background).

As we can see, there is a common attribute over all objects, the "Has Wheels" attribute, and because of that, this attribute is placed on the maximal. Moreover, two attributes are specific for its objects ("Has Helix" and "Has Wings" regarding the "Helicopter" and the "Plane", respectively). Note that both "Helicopter" and "Plane" inherit the "Has Wheels" attribute due to the order aspect (top to bottom; maximal to minimal). Looking at the concept node where the "Plane" object is placed, its intents are "Has Wheels" and "Has Wings", while its extent regards only the object "Plane". If we look from the maximal perspective, its intent regards only the "Has Wheels" attribute, while

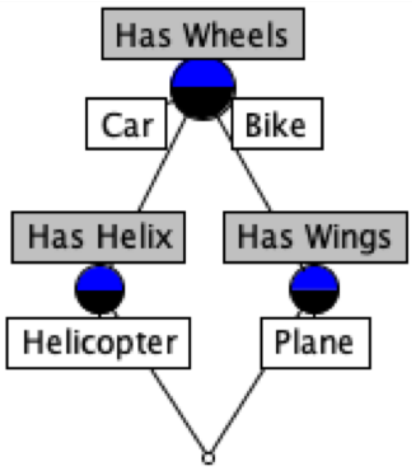


Figure 20 – Concept Lattice example

its extent includes all four objects of the context. Finally, the empty minimal shows no object has all of the attributes, and all of the attributes are linked to at least one object.

Next, based on the findings of Chapter 2 and some aspects presented in this Chapter, the Unification Explorer Framework is proposed.

4 The Unification Explorer Framework (UEF)

This Chapter is divided into five main parts:

1. The proposition of a framework (i.e., the UEF) to identify, in a set of models, the model that better represents the players' behavior (i.e., identifying a unified model of a considered context).
2. The UEF application to the two sets of models identified in Chapter 2.
3. The pros and cons of the proposed UEF.
4. UEF's final remarks.
5. The answering of the related RQs.

4.1 UEF Proposition

This framework is based on the holism idea (presented in Section 1), as it explores and sums each piece of knowledge (the parts and their interactions) from different models (a context) to identify a unified view of human behavior. The input models define the context scope. The only requirement to apply the UEF is the formatting of the considered models in the hierarchical structure depicted in Section 2.6.1. To reach the UEF objective, a set of steps (see Figure 21) must be followed to assess whether there is already a unified model that covers the characteristics of all the other models or not. The UEF provides a process for identifying a general[1] model in a specific context through one of two possible ways:

1. By promoting an existing model as a general one, or
2. By proposing a new general model based on all characteristics of the considered models.

- **Extraction Step**: consists of reading about all of the models and extracting all of their characteristics, naming each one as a "General Characteristic" (GC), to a separate list, named "General Characteristics List" (GCL). This process must follow the reliability guideline explained in Section 2.6.1 to avoid loss of information.

[1] A general model in this thesis means a model that contains all the characteristics of a given context (i.e., a set of models).

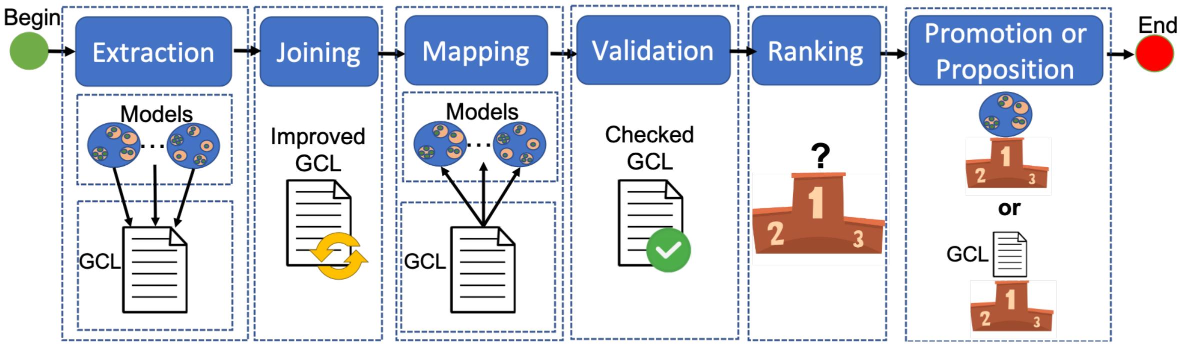


Figure 21 – UEF overview

Given the holism concept, since in this step all parts from the considered models are identified, and each one is viewed as a source of knowledge, this step regards the sum of the parts (i.e., the GCs) of a context.

- **Joining Step**: all of the GCs are analyzed and the similar ones are joined, also, enhancements are performed when one GC adds details to another. Next, general profiles (GPs) are proposed to group the GCs based on their similarities (i.e., when GCs approach the same subject, like social interactions), where each GC receives a code to facilitate its referencing in the "Mapping Step". At the end of the "Joining Step", the GCL is summarized and improved compared to its version in the "Extraction Step". Given the holism concept, this step regards the addition of the knowledge present in the interactions between the parts (i.e., the interactions between the GCs), meaning that the GCL corresponds now to the unified view of the context (the next steps will verify whether this view is contained inside the models or not).

- **Mapping Step**: the mapping process consists of linking each profile from the original models to the improved GCs from the previous step, providing a map that points to all of the GCs present in each profile from the original models and all of the sources for each GC (i.e., the models where the GC originated).

- **Validation Step**: uses the map from the previous step to verify for each profile from the original models if its linked GCs contains the original meanings of its descriptions, allowing the identification and correction of possible mistakes. This step is linked to the reliability guideline explained in Section 2.6.1 and can be seen as a double-check over the previous steps executions by ascertaining that no pieces of information were lost during the analysis, and the proposed linkages are correctly anchored.

- **Ranking Step**: the ranking process aims at verifying each model's coverage according to all GCs. This means that the number of GCs linked to each model would be counted, where the greater the number, the greater the coverage. Note that when a model has been compared to the GCL, it has been compared to all of the models concepts as the GCL contains all of these concepts.

- **Promotion or Proposition Step**: if, after ranking the models, there is one that contains all GCs of the GCL, this model will be promoted as a unified model, as it contains all of the characteristics present in the other models. However, if no model contains all of the characteristics, the GCL, together with its GPs, will be proposed as a general model (note that this composition already follows the proposed hierarchical structure of models). In conclusion, regardless of whether a model is promoted or a new one is proposed, a unified model will be identified as the final result of the UEF application.

An important note about this framework is that it does not weight models by their complexity, depth, or novelty degree, but by their coverage. It means that all models are considered equally important and the relevance of a model is analyzed from a quantitative perspective, the number of GCs. More details are presented in Section 4.3, where the pros and cons of the proposed framework are discussed.

Given the "Joining Step", it is essential to highlight that the resultant GC of an enhancement must always contain the descriptions of the original ones, with no exclusions. Assuming two GCs of the "Extraction Step" as "enjoys motivating others" and "enjoys teaching others", a researcher would propose to join them to enhance the abstract idea of social interactions by the following GC "enjoys teaching or motivating others". However, if the GC "enjoys motivating others" was instead "enjoys motivating others when they lose", the resultant GC should carry this conditional aspect of "when they lose", like "enjoys teaching, or when other players lose, enjoys motivating them".

To give an example of manual mistakes corrected by the "Validation Step", let us assume: a model with its original characteristics as "prefers painting as a relaxing activity" (Orig_1 for short) and "prefers driving slowly" (Orig_2), and enhanced GCs of the "Joining Step" as "prefers painting as a relaxing or fun activity" (Enhan_1) and "prefers driving fastly" (Enhan_2). Giving the linkages of the "Mapping Step" as Orig_1 with Enhan_2 and Orig_2 with Enhan_1, the assessment is ready to start. First, the original characteristics descriptions are confronted with the linked ones of the mapping. With this, it is possible to verify that Orig_1 and Orig_2 were wrongly anchored as their descriptions do not fit with the enhanced ones (i.e., painting is not driving and vice-versa), being the correct linkages: Orig_1 with Enhan_1 and Orig_2 with Enhan_2. Next, it is checked whether the original models' descriptions are contained inside the linked enhanced GCs or not. In this case, Orig_1 is ok since its description "prefers painting as a relaxing activity" is contained inside of "prefers painting as a relaxing or fun activity" (Enhan_1). However, it is not the case for Orig_2, as the idea of "slowly" is missing on Enhan_2; thus, to correct this mistake, the Enhan_2 is rewritten as "prefers driving slowly or fastly", solving the inconsistency.

Regarding the "Ranking Step", additional analysis can be done using the knowledge representation of Concept Lattices through their inherent aspects of maximal, minimal, intent, and extent (such as introduced in Section 3.8). For example, it is possible to identify in a lattice arrangement, the order of relevance of each GC, the coverage of each model, and the similarities between them.

Besides the capacity of Concept Lattices portraying similarities between objects through the sharing of intents, such identification can also be represented in terms of percentage. Applying it to the proposed hierarchical structure of models, Equation 4.1 considers the perspective from a model A toward a model B to compute the similarity between these models:

$$\text{Similarity} = \frac{AB_{sharedProfileCount}}{A_{ProfileCount}}. \tag{4.1}$$

Where $AB_{sharedProfileCount}$ regards the number of shared profiles between models A and B (i.e., the number of profiles that are linked to the same GCs), and $A_{ProfileCount}$ the number of profiles present on model A. Note that the similarity between A and B can be different compared to the similarity between B and A (by replacing $A_{ProfileCount}$ for $B_{ProfileCount}$ in Equation 4.1). This approach differs from the Concept Lattice one in two aspects: (1) it considers only two models at a time whereas the lattice approach considers all models at once, and (2), it attributes a similarity degree between 0% and 100%, whereas the lattice approach portrays this degree according to the shared intents. This Equation can be used to analyze the similarity of two specific models, if desired. In addition, this analysis can be extended to identify a similarity degree of the whole context by computing the mean value of all possible combinations of models, where the higher this value, the more similar the models are.

4.2 UEF Applications, Results, and Discussions

This section presents the UEF applications for the two sets of psychological models identified in Chapter 2, where one contains 46 models of players' behavior and the other 21 of general human behavior.

4.2.1 UEF Application to Players' Models

Next, each UEF step application, considering the identified players' models, is presented together with analysis and discussions.

4.2.1.1 Extraction and Joining Steps Applications

The “Extraction Step” identified a total of 548 GCs from the 46 players’ models, which were put in the GCL. In the “Joining Step”, all GCs were analyzed, and the equal ones were joined. Enhancements were performed when one GC added knowledge to another, resulting in a total of 80 improved GCs. In addition, 21 GPs were proposed to group similar GCs. The improved GCL is presented in Table 22.

Table 22 – Improved GCL (General Characteristics List) - Players’ models

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
Initial experience (InitExp)	InitExp1	Initial awareness, initial interest
	InitExp2	Initial try (testing the game), and initial progress in-game (high exploration, low exploitation)
Acceptance to play (AcceptPlay)	AcceptPlay1	After an initial experience (i.e., a progress in-game), the game is approved (considering its mechanisms, rules, and challenges) and played longer (a balance between exploration and exploitation) (the identification or maintenance of points of engagement)
	AcceptPlay2	After the end of interactions with a game due to negative historical, new points of engagement (e.g., new challenges or places to explore) are identified, and the game is played again
	AcceptPlay3	After interacting with points of disengagement, new points of engagement (e.g., new challenges or places to explore) are identified and maintained, keeping the player motivated although the negative historical

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	AcceptPlay4	The player is persistent by keeping playing (hard-working) to overcome game challenges regardless of the game difficulty
	AcceptPlay5	After finishing the game, the game is replayed in the same way as before or with other decisions or configurations
Lost motivation (LostMot)	LostMot1	Lack of motivation/new challenges
	LostMot2	Absence of friends
	LostMot3	Uncomfortable experiences (the identification or maintenance of points of disengagement)
	LostMot4	The end of interactions with a game due to the absence of motivations to play entailed by a negative historical (i.e., a continued occurrence of points of disengagement)
	LostMot5	To be forced to play by following a routine or degree of participation defined by others that surpass a personal degree of comfort (an impairment of the volunteer aspect of playing)
Accumulation of profit (AccuProfit)	AccuProfit1	The accumulation/gathering of items, friends, quests, riches, rewards, accesses, status, titles, points, levels, prizes, badges, money, and achievements due to progress in-game
Observer (Obs)	Obs1	Absence of control
	Obs2	Learning without interaction
	Obs3	Awareness of the surroundings
Environment exploration (EnviExplo)	EnviExplo1	The desire to discover things (game as uncharted territory, a source of information)
	EnviExplo2	The appreciation of lights, sounds, and colors present in fantasy or realistic environments
	EnviExplo3	Imagination to interpret the environment and innovation/creativity to find new ways to discover it

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	EnviExplo4	The observation of others' social interactions
Escapism (Esc)	Esc1	Avoidance of real-life problems through a relaxing and fun activity (leaving behind/forgetting concerns, people, and routine; coping with anger; a comfortable refuge; stress relief)
	Esc2	Protection against embarrassment/shame provided by anonymity when a new identity is assumed (escapism linked to real-life personal-identification)
	Esc3	Avoidance of real-life loneliness
Social interactions (SocialInt)	SocialInt1	Positive interaction (social enjoyment) between players or NPCs in a short, mid, or long-term (having confidence, trust, or belief) with friendship or sense of group (e.g., cooperation, teamwork, mediation of a group, collaboration, propositions, guidance, coaching, motivating, teaching, knowledge sharing, personal problems sharing, knowing others (social discovery), social network, making friends, support, assist, help, caretaking, passing the time together (talking or not), feedback, competition, matches, battles, championships, duels, combats, trading, gifting, and item sharing)

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	SocialInt2	Negative interaction (aggressive) between players or NPCs (e.g., competition, conflict, discussion, controversy, persuading, matches, battles, championships, duels, combats, protest, complaint, stealing, treason, betrayal, murder, social pressure, taking advantage, and interference in others' gameplay)
	SocialInt3	Interactions toward players with a similar or neutral appearance regarding own group in real or game-world
	SocialInt4	The wish to change social dispositions/hierarchy regarding voting, anarchy, or cultural influences
	SocialInt5	The wish to be in harmony with a personal group (from the real-world or game-world) by keeping playing with them according to the group common/expected interaction frequency (following social norms) or by reacting according to the group expectancy regarding a given situation
Thrilling (Thril)	Thril1	Excitement and relief associated with enjoyable gameplay
	Thril2	Tension release (i.e., the concept of catharsis)
	Thril3	Interest in dramatic situations (extrapolated behavior, easy fun, visceral impact)
	Thril4	Interest in terrifying, frightening, or violent challenges/situations
	Thril5	Interest in high-risk situations
Story awareness (StoAwar)	StoAwar1	Awareness, imagination, innovation, creativity, progression (e.g., by finishing quests), and appreciation regarding fantasy or realistic stories portrayed by the game plot

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	StoAwar2	The wish to know/discover/understand/unfold everything in a plot considering self-avatar influences or not (e.g., by finishing all quests) or create a new story (considering a game as an unfolding story) by role-playing, dressing-up, or playacting (following an avatar's appearance expected behavior)
Solving of problems (SolvProb)	SolvProb1	The study, proposition, and application of new ideas/strategies/approaches (through innovation, imagination, creativity, exploration, and efficiency aspects) to solve mental problems of specific or complex situations with satisfactory or optimal results
	SolvProb2	The study, proposition, and application of different uses (i.e., different combinations) of game mechanisms to improve performance (through tactical thought, innovation, imagination, creativity, and exploration; the seeking for efficiency)
	SolvProb3	The wish to gain power/strength, to be the best player, to have the highest score
	SolvProb4	Reasoning and proposition of goals, the identification of opportunities with good rewards linked to low effort
	SolvProb5	The chase or accomplishment of challenges/objectives (a game as an obstacle course; hard fun)
	SolvProb6	The mastering evolution process (the wish to learn and develop skills (improving coordination and concentration); the continuous learning about game mechanisms (from trivial to complex); learning based on self-mistakes; the facing of more difficult challenges)

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	SolvProb7	The pleasure of having the needed skill to solve a challenge (neither as easy nor as difficult; not trivial) or achieve an objective (e.g., winning, beating others)
Recognition (Rec)	Rec1	Approval or recognition from others due to possessions regarding profit, riches, power (social status), or status (derived from achievements)
	Rec2	Interest in showing self-status or achievements (e.g., top positions on leaderboards, or rare achievements)
	Rec3	The desire to be in evidence (exhibitionism) and express personal opinions or values
	Rec4	The recognition of the player's attachment to the game through the receipt of rewards
	Rec5	The wish to prove superiority toward challengers or to convince others that an adopted rational thought is the best option for a given situation
Autonomy (Auton)	Auton1	Customization (through imagination, creativity, and innovation)
	Auton2	The sense of control (modeling, operating, maintaining a planned work, making choices, and checking the response/efficiency of decisions made) over game mechanisms, avatars, or challenges
	Auton3	The sense of self (e.g., self-identification with an avatar considering self-beliefs)
	Auton4	Personal interpretation (giving meanings)
	Auton5	The freedom to build/create things in fantasy or realistic environments
Player mastery (PlayMaster)	PlayMaster1	Complete understanding and dominance over game mechanisms

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	PlayMaster2	Completion of all challenges
	PlayMaster3	A high degree of shared capabilities
	PlayMaster4	A medium degree of shared capabilities
	PlayMaster5	A low degree of shared capabilities (specialist, a particular knowledge)
	PlayMaster6	High trust in self-knowledge and self-capabilities linked to decision-making (low exploration degree and high exploitation degree)
	PlayMaster7	The possession of the highest score or time spent playing
Unsocial (UnSoc)	UnSoc1	The absence of social interactions
Serene (Ser)	Ser1	A calm and peaceful behavior (light touch)
	Ser2	Preference for no/low-risk situations (safe gameplay)
	Ser3	Avoidance of tough challenges through hard-working
Content consumption (ContCon)	ContCon1	Interest in endless game progress supported or not by a continuous content generation (i.e., new challenges) associated with new rewards
	ContCon2	A fair exchange between the player's effort and the obtained reward (cheatings not allowed)
	ContCon3	Interest in receiving rewards linked to defeating enemies or eliminating threats
	ContCon4	Interest in receiving rewards linked to delivering items
	ContCon5	Interest in receiving rewards linked to delivering messages
	ContCon6	Interest in receiving rewards linked to collecting items (peacefully or not)
	ContCon7	Interest in receiving rewards linked to escorting

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	ContCon8	No interest in receiving rewards linked to helping others
Security (Sec)	Sec1	The secrecy of the player's personal (real-world) data (anonymity)
	Sec2	Stable access to the game
Extension (Ext)	Ext1	According to personal in-game experiences, the perception of the game-world is extended to the real-world, influencing the player's self-identification or spatial comprehension (a partial joining between two worlds)
	Ext2	The perception of being spatially immersed in the game-world (an extension of the real-world spatial presence to the game-world spatial presence)
Playing (Playing)	Playing1	Playing as voluntarily interactions toward a game environment (based on chance/opportunity or randomness; encompassing learning and thinking), interest in playing (engagement to play derived from identified and maintained points of interest) as a need (obsessive, essential) or a desire (not obsessive, not essential)
	Playing2	Playing as a leisure/recreational activity that can improve mood (funny, challenging, interesting, exciting, entertaining, relaxing, and pleasurable)
	Playing3	Playing as a pastime/filling-time activity that refrains/alleviates boredom
	Playing4	A physical or mental experience that can be common in real-life (realistic experiences) or not (fantasy experiences; a hallucination of the real-world; unreality)

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
Burnout (Burn)	Burn1	Loss of opportunity to play

In addition to the improved GCL descriptions, the considered game context allowed the grouping of all GPs into General Topics (GTs) to highlight more abstract ideas, like "Players' Preferences" (i.e., desirable game settings), "Players' Motivations" (i.e., pleasurable activities), "Players' Status" (i.e., transitory states during a player lifetime in-game), and the "Essence of Gameplay" (i.e., abstract ideas about playing; the general descriptions of the game context applicable to all of its models). Table 23 links the relation of each GP to the aforementioned GTs.

Table 23 – General Topics (GTs) of each General Profile (GP) - Players' models

General Topic (GT)	**General Profiles (GPs)**
Players' Preferences	Autonomy, Content consumption, Observer, Serene, Security, Thrilling, and Unsocial
Players' Motivations	Accumulation of profit, Environment exploration, Escapism, Recognition, Social interactions, Solving of problems, and Story awareness
Players' Status	Acceptance to play, Burnout, Extension, Initial experience, Lost motivation, and Player mastery
Essence of Gameplay	Playing

4.2.1.2 Mapping, Validation, and Ranking Steps Applications

Moving to the "Mapping Step", the resulting map can be found in Appendix A. The "Validation Step" was applied using the generated map to ascertain that all links are correctly anchored and that all descriptions of the original models are contained inside the linked GCs. The same map was used in the "Ranking Step" to generate the Ranking depicted in Table 24. Note that it is impossible to have models with a GC count of zero as the GCL was built based on them. Moreover, each count portrays the degree each model contributed to the GCL regarding the coverage perspective.

Table 24 – Ranking result - Players' models

Ranking	**Model**	**GC count**
1st	Marczewski	29

To be continued

Ranking	Model	GC count
2nd	VandenBerghe	22
2nd	Olson	22
4th	Smith	21
4th	Laws	21
6th	Bateman and Boon	18
6th	Kellar et al.	18
6th	Yee	18
6th	Hunicke et al.	18
6th	Lin and Lin	18
11th	Brayshaw and Gordon	16
12th	Cook	15
12th	Zhu et al.	15
12th	Wu et al.	15
15th	Malone and Lepper	14
16th	Thue et al.	13
16th	Demetrovics et al.	13
16th	Eglesz et al.	13
19th	Bartle	12
19th	Sherry and Lucas	12
19th	Griffiths et al.	12
19th	Blacow	12
23rd	Weiller	11
23rd	Nacke et al.	11
23rd	Chou and Tsai	11
23rd	Frostling-Henningsson	11
23rd	Jansz et al.	11
23rd	Dickey	11
23rd	Haggis-Burridge	11
30th	Lazzaro	10
30th	Caillois	10

To be continued

Ranking	Model	GC count
30th	Reilly et al.	10
30th	O'Brien and Toms	10
34th	Koster	9
35th	Harbord and Dempster	8
35th	Jerčić	8
37th	Bulatov	7
37th	Cammarata et al.	7
37th	Koo	7
40th	Fairclough	6
40th	Pine and Gilmore	6
40th	Colwell	6
40th	Yee and Bailenson	6
40th	Hsu and Lu	6
45th	Kim and Ross	2
46th	Przybylski et al.	1

To better understand the relevance of each GC (considering their final form after the validation step) over all of the models, a Concept Lattice was proposed assuming M as the set of GCs and G as the set of models. It is important to highlight that the nearer an attribute is to the maximal, the more relevant it tends to be (i.e., the greater its extent tends to be), being the attribute on the maximal the most relevant one. According to the top part of this lattice arrangement presented in Figure 22 (the full lattice can be found in Appendix A; the red color was used to improve layout), Playing1 GC is present in all of the models as it is included in the lattice maximal, followed by the second and third most relevant GCs, Playing4 and Playing3 respectively. The Przybylski et al.(PRZYBYLSKI et al., 2009) model is placed on the maximal because it only has the Playing1 GC. In addition, there are 20 GCs that are specifics for their models: SocialInt3, Ext1, Ext2, Sec2, ContCon2, ContCon3, ContCon4, ContCon5, ContCon6, ContCon7, ContCon8, PlayMaster3, PlayMaster4, LostMot2, LostMot5, Burn1, Obs1, AcceptPlay5, Esc3, and Thril2. Looking from the maximal until the minimal perspective, the order where each GT first appears is: "Essence of Gameplay" with height 0 (the maximal), "Players' Motivations" and "Players' Preferences" with height 2, and "Players' Status" with height 3. This order portrays the authors' common points of view while modeling players' behavior, where all of them tend to deal with the "Essence of Gameplay" (as expected for this context), being the "Players' Motivations" and "Players' Preferences" more approached than the "Players' Status", as revealed by the height of each GT first occurrence.

Considering the analysis of coverage for each model, to build the Concept Lattice M was set as the link between models (i.e., when a model has the same GC as another

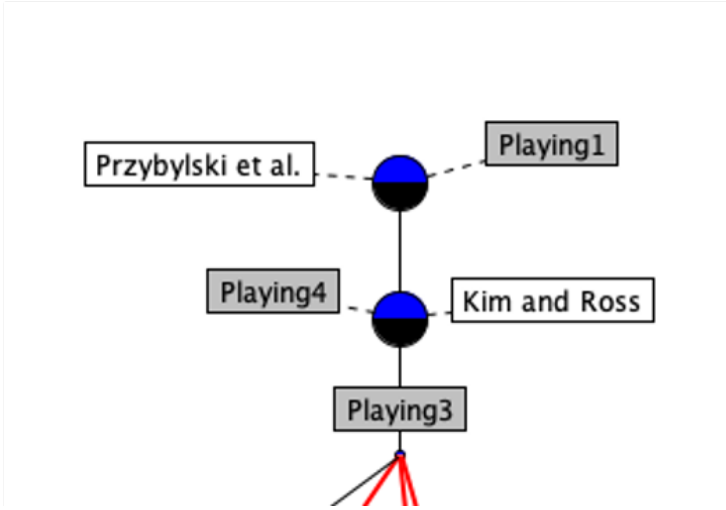


Figure 22 – GCs relevance (top vision) - Players' models

one) (by default, every model connects to itself) and G was the set of models. However, considering the existence of a general GC (as depicted by the maximal in Figure 22), the proposed linkage would entail a lattice with only one concept, as each model would link to all others under the context general descriptions. Bearing in mind this, the GCs that represent the "Essence of Gameplay" were removed together with the models that only contains such GCs (Przybylski et al.(PRZYBYLSKI et al., 2009) and Kim and Ross (KIM; ROSS, 2006)) to highlight the coverage of more detailed aspects. Figure 23 shows the top of this lattice until height 1 (the green circle indicates the concepts at height 1; the full lattice can be found in Appendix A).

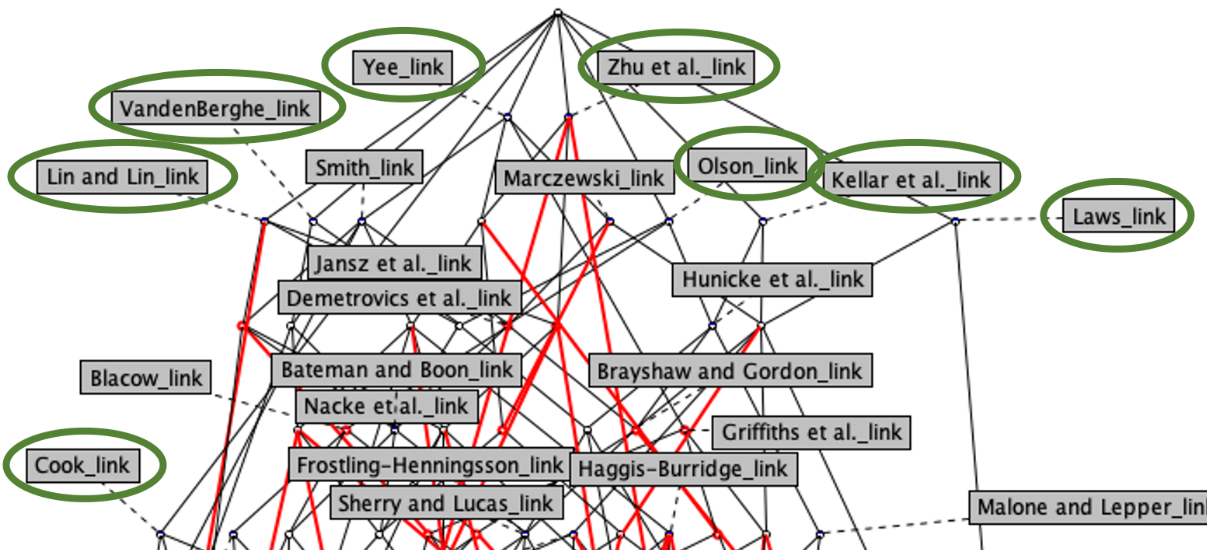


Figure 23 – Models coverage (top vision) - Players' models

As expected, the empty maximal confirms the "Ranking Step" result, as there is no model that covers all of the others. The empty minimal means that no model does not link to another, highlighting that no author had proposed a model containing only unique aspects of players. Comparing the Ranking presented in Table 24 with this lattice arrangement, one possible conclusion is that the Marczewski (MARCZEWSKI, 2015) model (1st place) is not as similar to the others because it is placed in the height 2 of the lattice. Meaning that, although it has the highest count of GCs, its GCs are not as common as the ones presented by the models of height 1. In addition, none of the models in height 1 share any GC, contrasting the spread points of view adopted by different authors. The lattice also depicts the existence of identical linkage between models by the sharing of two or more different objects in the same concept, such as: Jansz et al. (JANSZ; AVIS; VOSMEER, 2010) with Demetrovics et al. (DEMETROVICS et al., 2011), Bateman and Boon (BATEMAN; BOON, 2006) with Nacke et al. (NACKE; BATEMAN; MANDRYK, 2014), Frostling-Henningsson (FROSTLING-HENNINGSSON, 2009) with Sherry and Lucas (SHERRY et al., 2006), Colwell (COLWELL, 2007) with Hsu and Lu (HSU; LU, 2004) and Harbord and Dempster (HARBORD; DEMPSTER, 2019). Note that this comparison does not include their GC counts, but the models that are linked to them.

The similarity degree represented by Equation 4.1 was computed to all models, the results can be found in Appendix A in two perspectives, one considering the GT "Essence of Gameplay" (with a mean percentage value of 98.43%) and another disregarding it (with a mean percentage value of 31.58%). This value of 31.58% complements the findings of the lattice portrayed by Figure 23, by giving a percentage value associated to the dissimilarities of the models represented by the empty maximal and the presence of eight models at height 1.

4.2.1.3 Promotion or Proposition Step Application

Finally, the last step of the UEF was applied with consideration of the Ranking in Table 24. In this context, for a model to be promoted as a unified one, it should have a GC count of 80. However, as the 1st place model only has a GC count of 29, it was concluded that there is no unified model for this specific context. Thus, the GCL plus its GPs were proposed as a general psychological model of players.

4.2.2 UEF Application to Human-being's Models

Next, each UEF step application, considering the identified HBMs, is presented together with analysis and discussions.

4.2.2.1 Extraction and Joining Steps Applications

In the HBMs case, the "Extraction Step" identified a total of 1,766 GCs, which were put in the GCL, and the "Joining Step" application resulted in a set of 101 improved GCs. Moreover, 14 GPs were proposed to group similar GCs. The improved GCL is presented in Table 25.

Table 25 – Improved GCL (General Characteristics List) - Human-being's models

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
Long-term Psychological State (Motivational Source)	LtMotSour1	When a person's motivation comes from internal factors.
	LtMotSour2	When a person's motivation comes from external factors.
Long-term Psychological State (Environment Interpretation)	LtEnvInterpret1	A person that only accepts new information when it comes from real facts (expected or unexpected; desirable or not), avoiding uncertainty entailed by incomplete information.
	LtEnvInterpret2	A person that interprets new information based on abstract ideas through his/her imagination or creativity.
	LtEnvInterpret3	A personal automatic and immediate judgment of new situations as desirable or not, threatening or not, challenging or not, dangerous or not, gross or not.
	LtEnvInterpret4	When a person identifies the necessity to act in a new environment configuration (e.g., a new obstacle) to avoid an impairment of a human need.
	LtEnvInterpret5	The natural behavior of identifying ways of interaction toward the environment through curiosity, leaving aside efficiency aspects.
	LtEnvInterpret6	When a person assembles environment characteristics to compose a hedonistic or aesthetic view of life through smells, sounds, sights, tastes, and textures.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
Long-term Psychological State (Decision Making)	LtDecMak1	When a person uses logic (pros and cons) to make decisions based on his/her own perception or from others' ideas, being all pros and cons accepted as reliable.
	LtDecMak2	When a person makes decisions based on harmony toward others, following social and personal standards.
Long-term Psychological State (Environment Interaction)	LtEnvInteract1	When a person follows plans to keep the environment under control to maximize/maintain his/her pleasure or diminish/alleviate his/her pain, using or not of social interactions (e.g., impositions, persuasions, or cooperation) during the process.
	LtEnvInteract2	When a person prefers to adapt to the environment instead of controlling it; absence of control.
Short-term Psychological State (Negative)	StNeg1	When a person is uncertain about not desirable changes that can happen in the environment.
	StNeg2	When a person is certain about not desirable changes of the environment.
	StNeg3	When a person is upset after an undesirable interaction (expected or not) with an object, a person, an obstacle, or a situation that affected him/her goals (e.g., a failure), entailing in requests for support or the application of counter-measures; the impairment of human needs; lack of self-confidence. For example, when a wanted thing (tangible or figurative) is lost, taken away, there are chances to become scarce or it is more difficult to achieve due to a new obstacle.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	StNeg4	When a person's mind is less perfect (i.e., no optimal decision making) due to inability to deal with a new situation; when a person is careless, presenting implausible or pseudo-random behavior due to poor comprehension, confusion, or passive non-compliance.
	StNeg5	When a person is upset and searching for help due to an anticipation or identification of a threatening situation or a very difficult challenge according to his/her skills; the increased likelihood of some undesirable fact; the prospect of impairing a human need.
	StNeg6	When a person is upset due to the identification of an unattainable challenge (e.g., due to lack of abilities, miscomprehensions of the environment interaction rules, or the arising of a new obstacle).
	StNeg7	When a person is upset due to the occurrence of a fact presumed as undesirable for someone else who this person likes.
	StNeg8	When a person is upset due to the occurrence of a fact presumed as desirable for someone else who this person dislikes.
	StNeg9	When a person is upset due to the occurrence of an expected and undesirable event.
	StNeg10	When a person is upset due to the no occurrence of an expected and desirable event.
	StNeg11	When a person is disapproving his/her own action and its consequences that affected him/her or others (i.e., undesirable results).

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	StNeg12	When a person is apprehensive about his/her status or the consequences of his/her actions due to the possibility of receiving some punishment that would affect his/her relationships or status.
Short-term Psychological State (Positive)	StPos1	When a person is happy due to the occurrence of a fact presumed as desirable for someone else who this person likes.
	StPos2	When a person is happy due to the occurrence of a fact presumed as undesirable for someone else who this person dislikes.
	StPos3	When a person is happy due to the occurrence of an expected and desirable event.
	StPos4	When a person is happy due to the no occurrence of an expected and undesirable event.
	StPos5	When a person is admiring his/her own action and its consequences that affected him/her or others (i.e., desirable results; the pleasure entailed by a self-competence).
	StPos6	When a person is admiring someone else's action and its consequences that affected him/her or others (i.e., desirable results).
	StPos7	When a person desires to acquire things or be in situations of satisfaction; the wish to attain human needs.
	StPos8	When a person has happiness, contentment, or well being due to the occurrence of some event (caused by him/her or other; regardless if the way to get it done was painful) that affected him/herself, the possession of desirable objects, or social position in a group (a positive social relationship); self-confidence; the attainment of human needs (i.e., the achievement of objectives; the overcoming of obstacles).

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	StPos9	When a person is searching for new desirable things.
	StPos10	When a person is considering, understanding, and accepting others' opinions.
	StPos11	When a person is happy due to the increased likelihood of some desirable fact; the prospect of attaining a human need.
	StPos12	When a person is happy due to the identification of an attainable challenge.
Mid-term Psychological State (Positive)	MtPos1	When a person likes or appreciates another person or an object, liking or appreciating it even more as he/she/it is more explored and understood; a continued attainment of human needs.
Mid-term Psychological State (Negative)	MtNeg1	When a person has antagonism, counter-empathy, does not care, dislikes, or disapproves an object, another person's actions (and their consequences), a group behavior, or a situation (due to the continued impairment of some human need or the presence of elements that go against social and respectful standards).
	MtNeg2	When a person is resigned; when a person accepts something unpleasant as it is very difficult to change it.
Mid-term Psychological State (Neutral)	MtNeut1	When a person has never interacted with another person, an object, or a situation, presenting a neutral position regarding it/him/her (neither liking nor hating it).

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
Long-term Psychological State (general human nature)	LtGen1	When a person is cautious.
	LtGen2	When a person has perseverance (i.e., a hard-working and diligent person) to accomplish an objective, following a strict routine and working for many hours to overcome whatever is the obstacle, hardly giving up of it; to be persistent until the goals are achieved.
	LtGen3	When a person is responsible (altruist), taking into account others' perception.
	LtGen4	When a person has empathy/care toward others or objects, working together aiming shared goals through persuasion, understanding, leadership, support, motivation, guidance, the proposal of changes, solving of disagreements, the strengthening of relationships, and the boosting of abilities.
	LtGen5	When a person is sociable, appreciating the presence of friends and familiars and wishing to maintain and create new friendships; the seeking for positive social interactions.
	LtGen6	When a person is unsociable, not appreciating or caring about others.
	LtGen7	When a person is sincere (reliable and humble).
	LtGen8	When a person has enthusiasm, doing many leisure activities, enjoying of funny moments with a light-heart and no concerns.
	LtGen9	When a person desires to attain materialism needs.
	LtGen10	When a person desires to attain power needs.
	LtGen11	When a person desires to attain achievement needs; achiever.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	LtGen12	When a person desires to attain self-esteem needs.
	LtGen13	When a person desires to attain social needs.
	LtGen14	When a person desires to attain information needs.
	LtGen15	When a person desires to attain sensual needs.
	LtGen16	When a person desires to realize dreams.
	LtGen17	When a person desires to attain organic needs.
	LtGen18	When a person desires to attain safety needs.
	LtGen19	When a person is cordial and respectful.
	LtGen20	When a person is insecure and frequently needs to receive support from others, such as love, sympathy, protection, advice, and reassurance to feel safe.
	LtGen21	When a person is not reliable due to the continuous change of his/her own opinion or values depending on the circumstances.
	LtGen22	When a person is pessimist, uneasiness, self-belittling, moody, anxious, and easily annoyed (i.e., excitable; irritability), having a fragile emotional state (not desirable emotional states last longer and are more intense for both short-term and mid-term perspectives); Neuroticism.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	LtGen23	When a person has a stable emotional state (i.e., auto-control) with fair and honest behavior, being sober.
	LtGen24	When a person is passive.
	LtGen25	When a person is peaceful, avoiding any kind of conflict.
	LtGen26	When a person is quiet, reserved, discreet, or serene.
	LtGen27	When a person is open to or appreciates changes that take into account new facts and/or mutual benefits, avoiding routines.
	LtGen28	When a person is rigid, ignoring harmony and care toward others to achieve objectives good to him/herself and to others.
	LtGen29	When a person is less suscetible to anxiety (relaxed).
	LtGen30	When a person is resourceful (creative).
	LtGen31	When a person is optimistic.
	LtGen32	When a person is reactive to real facts or imaginary possibilities (i.e., always reacting to a new situation (expected or not); the opposite of passiveness), protecting him/herself from threats, harms, or criticism, taking an offensive position when needed with an unstable emotional state, being impulsive without deliberations.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	LtGen33	When a person is aggressive to achieve own objectives or impose something to others, appreciating combats, discussions, and submission from others.
	LtGen34	When a person is subservient, not a leader, accepting blame or criticism without retaliation even when it is not fair.
	LtGen35	When a person likes competitions.
	LtGen36	When a person cherishes perfectionism, appreciating and struggling to maintain high standard patterns or status; the desire to be competent.
	LtGen37	When a person appreciates the freedom (i.e., autonomy) to follow own way, rules, and standards, breaking or ignoring restraints of the environment if needed; lonely.
	LtGen38	When a person wishes (consciously or not) that others desire him/her or his/her status or riches through a presentation of an auto-confectioned image that describes him/her in desirable terms as well as his/her status and riches (regardless if they are accurate or not).
Short-term Psychological State Degrees	StDegree1	When a person had, or not, control over the environment during the occurrence of an event.
	StDegree2	When a person had, or not, a sense of pleasure during the occurrence of an event; the attainment or impairment of human needs.
	StDegree3	When a person was calm or alert during the occurrence of an event.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
Self Perception	SelfPer1	The perception of being able to perform an activity, assume a role, or complete a challenge (i.e., to overcome an obstacle not trivial, neither too easy nor too difficult) in an efficient way according to his/her skills; when a person considers him/herself as competent.
	SelfPer2	The perception of being able to perform an activity, assume a role, or complete a challenge (i.e., to overcome an obstacle not trivial, neither too easy nor too difficult) very easily, without too much effort.
	SelfPer3	The perception of not being able to perform an activity, assume a role, or complete a challenge in an efficient way according to his/her skills; when a person considers him/herself as incompetent.
	SelfPer4	The perception of having, or not, the freedom to make decisions and control the environment.
	SelfPer5	The perception of being connected, or not, to another person together with the acknowledgment of own value.
	SelfPer6	The perception of being able, or not, to comprehend the environment rules and possible interactions.
	SelfPer7	When a person can comprehend, or cannot, his/her status or actions' consequences as good or not toward others.
	SelfPer8	The perception of being able, or not, to recognize the effects of own emotions.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	SelfPer9	When a person is able, or not, to maintain or improve self-performances standards through the identification of opportunities.
	SelfPer10	The perception of being prompt, or not, to act.
Physiological Human Needs	PhyHN1	The attainment or impairment of organic needs, such as eating, drinking, sleeping, etc.
	PhyHN2	The attainment or impairment of safety needs, such as personal and financial security, and health.
Psychological Human Needs	PsyHM1	The attainment or impairment of social/affiliation needs, such as having friends, intimacy, and family, making new friends, being part of a group, helping others, seeking for protection and aid, being able to accept abasement and reject undesirable social interactions.
	PsyHM2	The attainment or impairment of esteem needs, such as being recognized, respected, approved, accepted, and valued by others (in high esteem), and having status and importance in society.
	PsyHM3	The achievement, or not, of personal dreams, such as mate acquisition, parenting, abilities usage, and goals.
	PsyHM4	The attainment or impairment of materialism needs, such as the gain of possessions, the construction of something, the arrangement of objects, and the retention of objects.

To be continued

General Profiles (GPs)	**GC Code**	**General Characteristics (GCs)**
	PsyHM5	The attainment or impairment of power needs, such as having the capacity to: attack or injure others, avoid blame or punishment, impose one's desires, revenge, maintain self-respect and pride in a high level, and control the environment.
	PsyHM6	The attainment or impairment of achievement needs, such as overcoming obstacles, resisting influence or coercion, avoiding pain, avoiding failure, and being able to claim for attention (being dramatic if needed).
	PsyHM7	The attainment or impairment of information needs, such as relating facts, analyzing experiences, exploring, understanding and explaining the environment, acquiring knowledge from many different areas to satisfy personal curiosity, verifying generalizations and synthesis of ideas through logical thought.
	PsyHM8	The attainment or impairment of sensual needs, such as relaxing with another person, enjoying sensuous expressions, and forming an erotic and deep relationship.

In the same way that happened to the players' models, the HBMs context allowed the grouping of all GPs into General Topics (GTs) to highlight more abstract ideas, like "Short-term psychological aspects" (i.e., emotions and human needs), "Mid-term psychological aspects" (i.e., sentiments and self-perceptions), and "Long-term psychological aspects" (i.e., personality traits). Table 26 links the relation of each GP to the aforementioned GTs.

4.2.3 Mapping, Validation, and Ranking Steps Applications

The resulting map of the "Mapping Step" can be found in Appendix A. The "Validation Step" ascertained that all links in the map are correctly anchored and that all descriptions of the original models are contained inside the linked GCs. According to the same map, the "Ranking Step" generated the Ranking depicted in Table 27.

Table 26 – General Topics (GTs) of each General Profile (GP) - Human-being's models

General Topic (GT)	**General Profiles (GPs)**
Short-term psychological aspects (emotions and human needs)	Short-term Psychological State (Negative), Short-term Psychological State (Positive), Short-term Psychological State Degrees, Physiological Human Needs, and Psychological Human Needs
Mid-term psychological aspects (sentiments and self-perceptions)	Mid-term Psychological State (Positive), Mid-term Psychological State (Negative), Mid-term Psychological State (Neutral), and Self Perception
Long-term psychological aspects (personality traits)	Long-term Psychological State (Motivational Source), Long-term Psychological State (Environment Interpretation), Long-term Psychological State (Decision Making), Long-term Psychological State (Environment Interaction), and Long-term Psychological State (general human nature)

Table 27 – Ranking result - Human-being's models

Ranking	**Model**	**GC count**
1st	Northrop	59
2nd	OCC	50
3rd	Deci and Ryan	41
3rd	Plutchik	41
5th	Goleman (The Big Eigth)	39
6th	Smith and Ellsworth	35
7th	Goldberg	32
8th	Hunt	31
9th	Ekman	30
9th	Nakamura and Csíkszentmihályi	30
11th	Frijda	29
12th	Toprac and Abdel-Meguid	26
13th	Murray	25

To be continued

Ranking	Model	GC count
13th	Eysenck	25
15th	Maslow	24
16th	Zillmann (Affective Disposition Theory)	18
17th	Myers et al.	16
17th	Goleman (Competences)	16
19th	Zillmann (Mood Theory)	11
20th	Mehrabian	4
21th	Huizinga	3

The same procedure applied to the players' models to explore the relevance of each GC over all of the models was applied to the HBMs (assuming M as the set of GCs and G as the set of models). According to the top part of this lattice arrangement presented in Figure 24 (the full lattice can be found in Appendix A), it is possible to notice that no GC is present on all of the models, given the empty maximal. Also, at height 1 (pointed out by green circles), the presence of seven different GCs highlights the distinct occurrences of them, namely, SelfPer6, StDegree1, PsyHM1, LtGen13, LtGen14, PsyHM7, and LtEnvInterpret5. In addition, there are 24 GCs that are specifics for their models: LtMotSour1, LtMotSour2, StNeg1, StNeg7, StNeg8, StNeg9, StNeg10, StPos1, StPos2, StPos3, StPos4, MtNeg2, LtGen19, LtGen21, LtGen24, LtGen25, LtGen28, LtGen29, LtGen31, LtGen34, LtGen35, LtGen38, SelfPer9, and SelfPer10. It means that different authors do not simultaneously approach these GCs, highlighting their different points of analysis of human behavior. Looking from the maximal until the minimal perspective, all GTs first appear at height 1. That shows that no topic is more approached than the other.

Turning to the analysis of coverage for each model, a Concept Lattice was built considering as M the links between models (i.e., when a model has the same GC as another one) (by default, every model connects to itself) and G as the set of models. Figure 25 shows this lattice arrangement.

As we can see, the lattice maximal contains 13 models, meaning that all models of the context share at least one GC with these models. Moreover, there are few cases where the models do not connect to all others, regarding the models: Mehrabian (MEHRABIAN, 1980; MEHRABIAN, 1995; MEHRABIAN, 1996; RUSSELL; MEHRABIAN, 1977), Goleman (Competences) (GOLEMAN, 1998), Huizinga (HUIZINGA, 2014), Zillmann (Mood Theory) (ZILLMANN, 1988; ZILLMANN, 2015), Zillmann (Affective Disposition Theory) (ZILLMANN, 1995; ZILLMANN, 1996), Maslow (MASLOW, 1968), Myers et al. (MYERS et al., 1998), and Eysenck (EYSENCK, 1967). Comparing the Ranking presented in Table 27 with this lattice arrangement, one possible conclusion is that the Northrop (NORTHROP, 1974; NORTHROP, 1984) model (1st place) is as similar to the others because it is placed in the maximal of the lattice together with other 12 models. Meaning

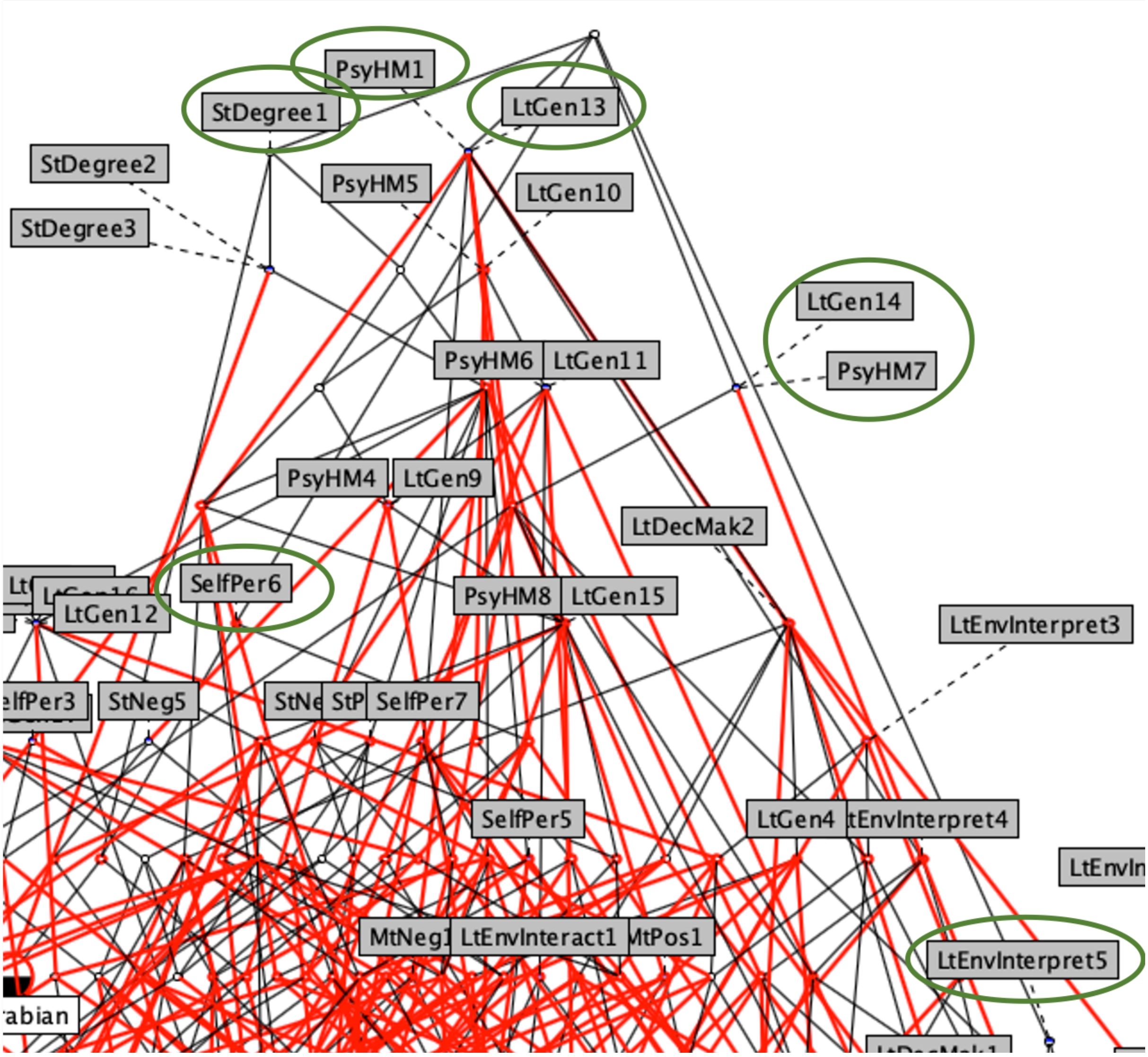


Figure 24 – GCs relevance (top vision) - Human-being's Models

that, even though it has the highest count of GCs, its GCs are as familiar as the ones presented by individual models.

The similarity degree represented by Equation 4.1 was computed to all models; the results can be found individually in Appendix A. The mean percentage value of the HBM context was 82.78%. This value complements the findings of the lattice portrayed by Figure 25, by giving a percentage value associated with the similarities of the models represented by the maximal intent that contains 13 models links from the total set of 21 models.

Interestingly, it was observed that psychological or physiological human needs notions were the basis for almost all of the 21 models. Where 20 considered the notions of psychological human needs, whereas 12 the ones of the physiological needs. This fact foments the proposed link between models depicted in Figure 7 in Section 2.6.3, where the psychological human needs play an essential rule to link long, mid and short-term

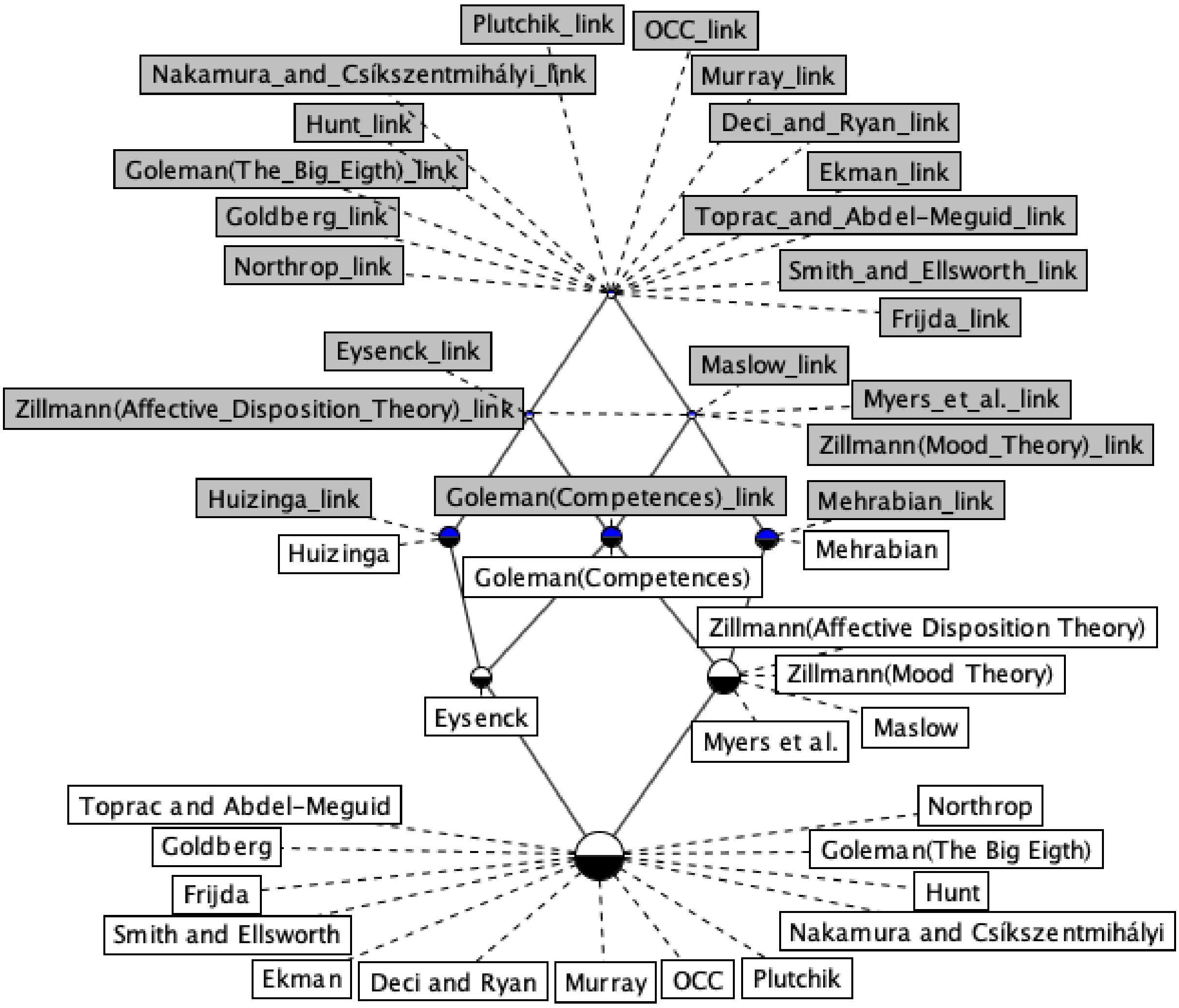


Figure 25 – Models coverage - Human-being's Models

psychological aspects.

4.2.4 Promotion or Proposition Step Application

The last step of the UEF application to the HBMs considered the Ranking in Table 27. In this context, for a model to be promoted as a unified one, it should have a GC count of 101. However, as the 1st place model only has a GC count of 59, it was concluded that there is no unified model for this specific context. Thus, the GCL plus its GPs were proposed as a general psychological model of human behavior.

4.3 UEF's Pros and Cons

This section presents the pros and cons of the UEF design. Also, the pros and cons of using a specific model instead of a unified one are approached in three different scenarios regarding the game context. The portrayed findings suggest that the unification of models is worth as a unified model provides more knowledge to support the interpretation and understanding of human behavior.

4.3.1 UEF's Design Pros and Cons

Next, the consequences of applying the adopted UEF's design over psychological models are presented considering limitations, benefits, counterpoints, and justifications.

- Cons:

  The UEF demands inputs that follow the idea of hierarchical structure, depicted in Section 2.6.1, obliging the formatting of psychological descriptions into concepts of "Model", "Profile", and "Characteristic". Even though most of the identified models already followed such composition, other models may present a more complex hierarchy that could not fit well with the proposed one, entailing an improper consideration of other authors' findings where pieces of information could not be appropriately considered.

  The framework is susceptible to personal bias. Although a reliability guideline is proposed together with a validation step to mitigate the bias, it is possible that considering the same input models, different researchers obtain dissimilar results. We understand that it is a natural aspect of the proposition of psychological descriptions since such pieces of knowledge arise from persons' cognition in interpreting environments (i.e., interpreting other persons in this case), which is essentially varied due to individuals' intellectual functions, such as reasoning, evaluation, judgment, and memory (ECKARDT, 1995). We also understand that this nature is responsible for the crescent universe of models, where, even more, different points of view arise to express a part of the complex essence of human beings.

  It is possible to interpret the proposed framework as having a reductionist component since descriptions of complex theoretical phenomena from varied sources can be broken into smaller parts, named in this thesis as "Characteristics". However, such breaking is adopted to identify atomic essences between different models, allowing the recognition of similarities between them.

  The relevance of models is quantitatively weighted using their coverage (number of GCs), disregarding any qualitative aspect of complexity, depth, or novelty degree. It implies that models usually adopted by academia or industry due to their qualitative aspects can have an unexpected position in the proposed Ranking due to specificities that do not encompass a substantial coverage. Nevertheless, as the final result of the proposed framework is a unified model of a context, even if a commonly used model is labeled as not general, its descriptions will be present in the proposed unified one, which also inherits its qualitative aspects.

- Pros:

  The lack of appreciation over the manner that authors name their propositions, such as theories, models, or archetypes, allows the proposed framework to explore and identify similarities and complements of different approaches by mapping and pointing common aspects, as showing the ones that are not so approached. As examples, it is possible to cite the social and disengagement aspects, where the social one is a common topic, whereas the other is almost not considered by the 46 identified players' models.

  A resultant unified model can be updated over time by reapplying the framework considering this model and uncovered ones.

  The definition of a unified model is founded on the holism concept, which carries the idea of defining a whole as the sum of the parts and the parts' interactions capabilities. Such aspect allows the enhancements of pieces of knowledge using varied sources.

  Researchers can apply the proposed framework using other sets of psychological models from different contexts (e.g., economics) to verify if their adopted models have the quality of a general model, or just identify and utilize the resultant unified model.

  The sharings of GCs by different models can be used to verify the similarity degree between them, as well as a context overall similarity, such as proposed by Equation 4.1. This kind of measure allows the identification of how diverse are the current authors' descriptions of human beings' behaviors considering a given context (e.g., gaming, economics, philosophy, neuroscience, or artificial intelligence).

  The adoption of Concept Lattices allows exploring specific aspects of psychological models by assuming as objects and attributes desired features. The relevance of such features can be verified by identifying the sharing of concepts, the concepts height, and the presence of objects or attributes on the maximal and minimal. Also, the intent and extent of a concept allow identifying a feature influence over the whole lattice.

4.3.2 Unified Model Pros and Cons

This section discusses the pros and cons of adopting a unified model instead of a specific one. Using the proposed unified model of players as a reference, three scenarios regarding Game Analytics, community management, and player simulation are approached to provide different analysis perspectives. For all cases, the proposed unified model of players is compared to the specific one presented in Table 28.

Table 28 – Bateman and Boon model (BATEMAN; BOON, 2006) in a hierarchical table format

Model	Profile	Characteristics
Bateman and Boon	Conqueror	Achievements of all game challenges and recognition
	Manager	Solving of problems, proposition of strategies, and seeking to develop skills
	Wanderer	Fun experience attached to escapism (i.e., leave behind concerns of the daily life)
	Participant	Positive social interactions as a member of a group

4.3.2.1 Game Analytics Scenario

Let us assume a situation of churn management of a given game, where a researcher intends to enhance his/her prediction model by adding psychological features. Having, on the one hand, the Bateman and Boon model (BATEMAN; BOON, 2006), considered as a specific model for the hypothetical game, and on the other hand, the proposed unified model of players. Also, let us assume that this game logs telemetry data of players regarding the following 12 possible actions/events (here, an action is something that a player did in-game, whereas an event is something that affected the player in-game, being this player the origin or not of the event):

1. "Kill monster": when a player defeats an enemy.
2. "Die": when a player is defeated.
3. "Buy item": when a player acquires a new item.
4. "Sell item": when a player sells an item to another player or an NPC (non-player character).
5. "Acquire quest": when a player receives a new mission.
6. "Finish quest": when a player completes a mission.
7. "Join party": when a player becomes a new member of a group of players.
8. "Leave party": when a player leaves a group of players.
9. "Level up": when a player increases the level of his/her avatar (improving its strength).
10. "First dungeon entry": when a player enters a dungeon for the first time.
11. "Dungeon completed": when a player completes all challenges of a dungeon.

12. "Weapon upgrade": when a player improves his/her weapon with better attributes.

Next, a mapping can be done to verify for each possible action/event, which are the associated profiles. A possible assignment for both models are presented in Table 29. Given that psychological features could be generated, for example, one could count for each player, the number of occurrences of each profile, and use it as additional features to predict churn.

Table 29 – Example of mapping actions/events into psychological profiles

Action/Event	**Bateman and Boon Profile**	**Unified model profile**
Kill monster	Conqueror	SolvProb5, SolvProb7
Die	-None-	LostMot3
Buy item	-None-	AccuProfit1
Sell Item	Participant	SocialInt1
Acquire quest	-None-	EnviExplo1, StoAwar2
Finish quest	Conqueror, Manager	StoAwar1, SolvProb5, SolvProb7
Join party	Participant	SocialInt1
Leave party	-None-	SocialInt2
Level up	Manager	SolvProb3, AccuProfit1
First dungeon entry	-None-	EnviExplo1
Dungeon completed	Conqueror, Manager	SolvProb5, SolvProb7, PlayMaster2
Weapon upgrade	Manager	Auton5, SolvProb3, AccuProfit1

Note that other psychological aspects could also be associated with players' behavior by considering the frequency of actions and events that happens to each player. For example, one could link the profile "Unsocial" from the proposed unified model to players that have never searched for social interactions (assuming a boolean feature), or the profile "Serene" to the players that have a long sequence of "level up" events before starting to accomplish quests and completing dungeons (also a boolean feature).

The advantage of translating actions/events to psychological features is that it allows grouping different inputs to the same psychological essence, and given a historical sequence, the assignment of complex essences can also be performed. Furthermore, even though different works can suggest different features based on the same game and adopted model, the best configuration can be assessed by traditional Machine Learning metrics applied to the classifiers' final results, such as the accuracy, recall, precision, or the F1 score (as presented in Subsection 3.6.3.1).

It is essential to highlight that the proposed Game Analytics scenario is a simple example of how psychological features can be built based on psychological models and usage data, and other approaches that vary in complexity can be found in the state-of-the-art (such as depicted in Section 1 and Subsection 2.7). The encompassing of even more psychological aspects in the Game Analytics feature engineering process is a current trend that has presented relevant improvements, as the new features entail in better divisions of the hyperplane by prediction models. This fact emphasizes the need to identify means to use the existing psychological models better, being the UEF an option to do it by providing, as much as possible, descriptions to be attached to actions or events present on usage data.

Taking into account the proposed scenario, the pros and cons of adopting a unified model in the Game Analytics field are:

- Cons:

  The time spent assigning profiles to actions/events when a specific model is adopted is shorter than the time spent when a unified model is considered. According to the proposed example, while the specific model demanded for each one of the 12 actions/events four assessments, the unified one demanded 80 assessments per action/event. Depending on the number of actions/events present on a game log and the available time to enhance a prediction model, the use of a unified model could not be feasible.

- Pros:

  The unified model has a more significant potential to assign profiles to the action/event list of a game than the specific model. As shown in Table 29, while the specific model encompassed 7 of the 12 actions/events, the unified model assigned all of them. Moreover, the unified model allowed the assignment of 11 distinct profiles, which is greater than the number of profiles in the specific model (4). Such a higher degree of detailing suggests that the unified model is more accurate in pointing psychological aspects of players in comparison to the specific model, which is a desirable aspect when individualized analysis of players is a need.

  The adoption of psychological features supports a better interpretation of the internal rules created by prediction models compared to approaches that do not use them, since such features are explained by the psychological models that originated them. Such quality is the desired aspect of a growing tendency in the industry, named Explainable AI (XAI), which big companies, like Google[2], are interested. The idea behind XAI is increasing the transparency of prediction models by identifying the relevance of each feature and explain it. This fact implies that the more psychological

[2] For more details, please access this link: https://cloud.google.com/explainable-ai

features are present in a model, the better, since more pieces of information can be interpreted by XAI frameworks. Therefore, adopting a unified model is preferred over a specific one because it provides more detailed and varied descriptions, such as shown in the proposed example.

4.3.2.2 Community Management Scenario

The management of communities is another way to assess the players' feelings about a game, besides the Game Analytics one. Since complaints on social networks can be identified, and counter-measures can be applied (KOZLOV, 2018). However, such kind of approach is reactive. The better option is to be proactive and avoid players' disengagement by anticipating what they like most and then use it to "feed" them with desirable experiences (FONG, 2019). To do it, questionnaires can be applied to identify what the players like and dislike, guiding new content development.

Assuming that a new questionnaire will be written to identify the players' motivations of a given game, this writing process can be guided by the descriptions of a psychological model. Groups of questions can be proposed to represent the characteristics present on the model, merely demanding a Yes or No answer, having for each model's profile at least one question to represent it.

By considering the chase in identifying players' motivations, the adopted specific model fits well, since it portrays only motivational aspects. By contrast, the unified model demands some exclusions, since it depicts, besides players' motivations, players' preferences, status, and the essence of gameplay (i.e., the unified model's GTs present on Table 23). After selecting the profiles under the GT of players' motivations, the unified model is ready to be used, offering 27 profiles in comparison to the four of the specific model. Looking from the specific model perspective, the following five questions would be proposed to represent its profiles:

1. Would you like to receive new challenges when you finish all of the available ones? (Conqueror)
2. Would you like to have your score shown in a ranking visible to all players? (Conqueror)
3. Is of your interest receiving new mechanisms regarding the gameplay? Like new ways of locomotion? (Manager)
4. Do you play as a manner to forget the problems of daily life? (Wanderer)
5. Do you enjoy social interactions? (Participant)

Following the same idea, at least 27 questions would be proposed to represent the unified model's point of view (not exemplified in this thesis), where some of them can add details to the proposed ones of the specific model, whereas others can cover not approached aspects (such as when negative interactions between players is a desire, which is described by the SocialInt2 profile). Using the motivation linked to recognition as an example, while the specific model questionnaire approaches it in one question by asking if a player would like to have his/her score visible to all players in a ranking, the unified model provides at least five questions (regarding the profiles Rec1, Rec2, Rec3, Rec4, and Rec5) to approach the same topic, such as the following ones:

1. Would you like to receive the approval or recognition from other players regarding possessions like status, power, profit, or riches? (Rec1)
2. Do you have an interest in divulging your status or achievements on leaderboards visible to all players? (Rec2)
3. Would you like to have the opportunity to express your opinion or values in-game widely? (Rec3)
4. Would you like to receive rewards linked to your attachment to the game? Like receiving a gift for every year played. (Rec4)
5. Would you like to have a simulation environment to prove that your strategy is the best way to a given challenge? (Rec5)

Based on the players' answers, it is possible to guide the development of new content that fits the players' motivations. For example, given the answers to the five recognition related questions, a fan pointing system, an achievements' leaderboard, a special chat platform, a campaign to reward players periodically, and a simulation environment can be built to "feed" players with known desirable content. By contrast, if the specific perspective was considered, only the leaderboard would be implemented to "feed" the recognition desire of players.

Given the proposed examples, the pros and cons of adopting a unified model instead of a specific one in the community management scenario are:

- Cons:

  As the unified model tends to have more descriptions than the specific one, it would entail a more extensive questionnaire, whose answering may demand a more significant effort from the players than they are willing to spend. However, the answering of such a questionnaire would entail in more detailed descriptions of players' behaviors.

Depending on the questionnaire objective, the unified model must be reduced to consider only the profiles linked to this objective.

- Pros:

  A chosen specific model may not contain the descriptions of a desired objective (e.g., the identification of the current transitory state of players), whereas the unified one has a greater chance of having them due to its greater coverage of concepts.

  The unified model can improve a questionnaire built based on a specific model since questions can be reformulated to encompass a greater range of details. Such as happened in the recognition example.

  The GTs of the unified model can be adopted as guidelines to highlight what kinds of questionnaires should be used in a game. By writing a questionnaire for each GT, different players' assessments would be gauged regarding their motivations, preferences, transitory states, and the most prominent aspects that make them play games.

  The development of new content tends to be more assertive when it is based on the answers of a unified model based questionnaire rather than if it was based on the answers of a specific model based one, since more details are provided in the unified perspective about what players like and dislike.

4.3.2.3 Player Simulation Scenario

The third and last perspective to analyze regards player simulation. The idea of simulating players aims at truthfully performing human-like behavior in-game by making an avatar act and react believably through the use of algorithms, an idea usually denoted by the term "believable bot" (KERSJES; SPRONCK, 2016). This kind of approach differs from traditional AI algorithms, where the winning efficiency is the key aspect to follow (IGAMI, 2020). Here, the approach is similar to the concept of a "Turing Test" (TURING, 2009), but instead of verifying if the machine can be intelligent, the idea is assessing the degree to which the machine can act in the same way as humans do from a psychological perspective (HINGSTON, 2009b). It is interesting to highlight that some contests aim to judge the most believable bots, such as the 2K BotPrize (HINGSTON, 2009a).

Another appealing fact surrounding the simulation of players is the possibility to hypothesize that, in the same way that players enjoy playing with other players, this same kind of motivation could also arise, to a certain degree, from interactions with believable bots. Where in the best situation, a player would not be able to distinguish between a real player and a bot. With such quality in hands, game producers could use it to keep players motivated by providing "artificial friends" to them, as it is known that playing with friends is one of the last motivators for players to continue playing, although there are disengagement aspects (ZHU; LI; ZHAO, 2010).

Assuming the task of building a believable bot, psychological models can provide hints in how to ground perceived events and performed actions on a human-like behavior. Based on models' descriptions, it is possible to identify what a player likes, prefers, and the motivational stages, giving also details of what motivates and demotivates the players. In view of the adopted specific model, it is possible to propose an avatar behavior encompassing 5 motivational properties, such as follows:

- Motivations:

  The wish to accomplishing challenges

  Enjoy recognition

  The wish to discover new and better solutions to problems

  The detachment to routines of daily life

  The willing to be with others

It is essential to highlight that those properties must be confronted with the game environment that defines the possible sets of actions and events to allow a proposed algorithm to translate the psychological properties into actions in-game. Looking at the proposed unified model, it is possible to segregate its properties based on the GTs profiles. The quantity of properties per GT are as follows:

- Players' Motivation: 27
- Players' Preferences: 27
- Players' Status: 22 (Transitory motivational states)
- Playing GT: 4 (Key essences of playing)

Given the aforementioned descriptions, the pros and cons of using a unified model instead of a specific one in the player simulation scenario are:

- Cons:

  None.

- Pros:

  As human behavior simulation encompasses a varied range of possibilities, it is assumed that the more psychological properties are implemented in the algorithm, the better. Thus, adopting a unified model is preferred instead of a specific one, as more features are addressed. As an example, while the specific model pointed to five properties, the unified one illustrated 80.

4.4 UEF's Final Remarks

Given the pros and cons of the UEF's design and the adoption of a unified model, the following implicit hypothesis "the use of a unified model knowledge provides better insights to interpret and understand players rather than the use of a specific model, which can have less detailed descriptions" is assessed as accurate, supporting that unification can be worthwhile. Bearing in mind the design perspective, for each set of psychological models, the framework allowed the mapping and joining of varied psychological descriptions to provide an enhanced view compared to the views presented by individual models. As exemplified by the three proposed scenarios in the game context and extending these findings to the general human behavior context, unified views support better insights to interpret and understand humans' behaviors by offering more varied psychological descriptions, which can have different applications depending on the adopted background. Such support was exemplified as worth in the Game Analytics, community management, and player simulation scenarios, but other research areas can also take advantage of using a unified model since it possesses textual descriptions of human behaviors. In special, we highlight that the enhanced knowledge present in the proposed unified models comes from both academic and industrial approaches, which nurtures the models with the quality aspects of being holistically founded on the academic state-of-the-art and the professional knowledge of the competitive gaming industry.

It is essential to highlight that a UEF result is influenced by the quality of its input models (i.e., the given context). However, regardless of the models' degree of richness, the UEF will always give a final result based on the unified view of the specified context. Although some psychological models may not be considered in an execution of the UEF in a given context, the proposed structure composed by the GCL, GPs, and GTs allows the assessment of new models' impact. In this case, considering the addition of a new model, three possible results may be obtained: (1) the new model does not add new GCs, and consequently, the current general model is kept the same; (2) the new model adds new GCs, enhancing the knowledge of the general model; or (3) the new model is promoted as the general model. In sum, regardless of whether a current general model is out of date, the framework can always be reapplied to consider a new model, resulting in the identification of a general model, which may or may not be an improved version of its predecessor.

Considering the crescent universe of models, a limitation of this thesis regards the identified ones, as it is not possible to ascertain that all models applicable to games that exist are considered due to issues such as the possible secrecy of game companies. Thus the proposed general models can be seen as summarized views of players' behavior, but not as comprehensive ones. However, the proposed framework has the potential to produce such aspects. The framework is also susceptible to bias, which is mitigated by the proposed reliability guideline and validation step. All the information is available for future replication if desired.

4.5 UEF's Related RQs Answers

This section contains the answers to the six RQs related to the Unification Explorer Framework proposition and obtained results.

4.5.1 Answer to RQ4

The answer for the research question "*Is it possible to link a profile of one model to the profile of another model?*" is yes, it is possible. It is exemplified by the UEF's "Joining step", where the atomic essences of each profile of specific models, the GCs, can be linked to similar or complementary GCs of other models, which are stored in the GCL.

4.5.2 Answer to RQ5

To answer this question ("*Can psychological models be ranked?*") it was proposed to rank models from a quantitative perspective, where the greater the number of GCs a model has, the greater is its coverage. However, this thesis does not approach any qualitative measure to rank models by considering their degree of complexity, depth, or novelty. Therefore, the answer to the RQ5 is "From a quantitative perspective, models can be ranked according to their GC counts; however, they cannot be ranked from a qualitative perspective due to a lack of procedures to qualifies psychological models based on their complexity, depth, or novelty degree".

4.5.3 Answer to RQ6

The answer for research question 6, "*Is it possible to combine models?*", is yes, it is possible. It is precisely what the UEF does, as it combines the psychological descriptions of different models to identify or propose a general point of view about human behavior. An interesting fact of the state-of-the-art is that, currently, the identification of a general model is an open problem (SCHRÖDER, 2004; LOIZOU et al., 2012); however, the solution to this problem can now be chased by using the UEF, as it provides means to going toward a general view of human behavior through the continuous addition of uncovered models.

If the proposition of new psychological models stops in the future, it will be possible to generate a final unified model.

4.5.4 Answer to RQ7

Even though two different sets of psychological models were identified in the proposed SLR, and each one of them had its models ranked following the UEF's "Ranking step" procedure (as portrayed in Tables 24 and 27), no psychological model was promoted as a unified one (i.e., no model had a GC count equal to its GCL GC count). Therefore, our answer for the research question "*Is there a general psychological model that can portray all the players' aspects?*" is "According to the set of identified models, and excluding the models generated by the UEF, no, there is no model that can portray all the players' aspects, as each one of the identified models considers a set of concepts (with low or high abstraction) that may not fit all other concepts from other models".

4.5.5 Final answer to RQ8

The initial answer presented in Chapter 2 for the research question 8, "*Are all models applicable to all game genres?*" was "No, they are not". However, by considering the possibility of obtaining a general model from the UEF, it is expected that the use of a general model encompasses a more significant number of game genres compared to the adoption of a specific one. Given this, by assuming that a general model was built based on at least one psychological model of each game genre, the final answer to the RQ8 is "No, they are not. The only model applicable to all game genres is a unified one built based on at least one model from each existent game genre".

4.5.6 Complementary answer to RQ9

In addition to the initial answer to research question 9, *"What are the advantages and disadvantages of using psychological models?"* presented in Chapter 2, we add all the seven pros portrayed by the Subsection 4.3.2, which highlights the benefits of adopting a unified model instead of a specific one. Given this, the advantages of adopting psychological models can be summarized as the nurturing of a better players' behavior comprehension, which can be linked to the mitigation of risky situations (e.g., churn), the development of more assertive content (e.g., through community management), and the propositions of more believable NPCs.

5 Proposed Method

This Chapter aims at answering the RQ 10 “To what extent characteristics of usage data can be used to identify psychological profiles?”, RQ11 “How an identified profile on usage data can be assessed?”, and describing additional advantages and disadvantages linked to the RQ9 “What are the advantages and disadvantages of using psychological models?”. As a starting point to develop a method that identifies players’ psychological profiles, first, it is needed to define this resultant profile, highlighting the required inputs and the psychological essences present on it. To do such a definition, the two identified general models are approached as reference candidates, where one of them is chosen according to two desirable aspects: simplicity and coverage.

On the one hand, the simplicity idea is linked to the chase of an input that requires, as less as possible, pieces of information from usage data, aiming at increasing the usability of the resultant profile to many kinds of games as possible. On the other hand, the coverage idea regards the encompassing of the more significant number of psychological essences to be added to the proposed profile, which are obtained from the input. Therefore, the proposed psychological profile is built based on the reference model that presents better trade between simplicity and coverage.

This Chapter is organized as follows: the composition of the proposed psychological profile of players is presented in Section 5.1 taking into account the two identified general models, the method assumptions are presented in Section 5.2 (highlighting additional aspects regarding RQ9), the method overview and its steps are described in Section 5.3, special considerations are pointed out in Section 5.4, the method assessment regarding the literature support and specialists analysis is discussed in Section 5.5, and the answers for RQs 10 and 11 are given in Section 5.6.

5.1 Proposed Psychological Profile Composition

The definition of this thesis’s psychological profile considers an analysis of the two identified general models, where based on their simplicity and coverage aspects, one is chosen to serve as the reference to extract psychological aspects from data. It is essential to remember that a psychological aspect of the proposed profile regards a feature, which can refer to one or more GCs occurrences of the chosen model, meaning that a single aspect of the proposed psychological profile of players can refer to multiple GCs (as shown further in this Section). Besides, it is also important to bear in mind that each GC carries its own identification challenges entailed by its textual description complexity. In an ideal situation, all the GCs of the chosen model would be considered by features extracted from data according to a straightforward input (e.g., demanding only the players' actions[1]). In sum, to identify the simplicity and coverage aspects of each one of the considered models, their GCs must be analyzed from two points of view. One where the required pieces of information and their sources are identified, pointing to a simplicity degree, and the other where the needed procedures to identify the GCs occurrences on data are approached, highlighting a complexity degree attached to the coverage aspect. By summing up all the findings for each model, a final decision can be taken.

Beginning with the unified player model perspective, it is possible to see that it presents GCs that vary in complexity, which entails different approaches to identifying such GCs occurrences on usage data. For example, from simple ones like LostMot2 (absence of friends) where it is just needed to check if a player had stopped his/her social interactions, to complex ones like SolvProb2 (the study, proposition, and application of different uses (i.e., different combinations) of game mechanisms to improve performance (through tactical thought, innovation, imagination, creativity, and exploration; the seeking for efficiency)), which requires deep notions of when a player is trying something new to improve performance or not. Given the varied conditionalities present on the GCs of this model, it was observed that each one requires a specific identification rule, which is sometimes similar to another one. It means that 80 different procedures are required to identify all the 80 GCs of this model. According to this initial analysis, the needed input to identify all of these GCs regards usage data (containing the players' actions) and questionnaires, as there are aspects like InitExp1 (initial awareness, initial interest) that cannot be identified on data.

Moving to the unified human-being model perspective, it also presents GCs that vary in complexity. Like the simple LtGen6 (when a person is unsociable, not appreciating or caring about others), which its occurrence on data can be identified by checking when a person has never searched for social interactions, and more complex ones, like the LtDecMak2 (when a person makes decisions based on harmony toward others, following social and personal standards), which requires the acknowledgment of when a person is doing something to be in harmony towards other's desires. Also, it presents GCs that cannot be identified on usage data, like LtMotSour2 (when a person's motivation comes from external factors). Thus, the required inputs to identify all the 101 GCs are usage data (containing players' actions and the reference to the players affected by those actions) and questionnaires.

It is essential to remember that this thesis objective regards applying a non-invasive approach to consider a more significant number of players. Therefore, it is impossible to identify the occurrences of all the chosen model's GCs, regardless of whether it is the

[1] A complete description of a player's action must encompass the action/event name, the player ID, and the timestamp when the action/event happened.

unified player model or the unified human-being model because more than usage data would be needed. Given this, the proposed profile focuses only on the GCs, from the chosen model, that are extractable from usage data.

By comparing the two unified models perspectives, it was identified that the human-being one presents a greater coverage of concepts compared to the player one, given that it has more GCs (101 compared to 80) and the player model is inherently linked to the human-being model (as players are humans). However, it was also identified that such greater coverage demands more complex identification rules to retrieve each GC's individual occurrences, entailing a more complex input in some cases. Interestingly, the unified human-being model presents a property that we named "Aggregated psychological essences", which, if successfully implemented in a method, has the potential to simplify its GCs identification rules. This property is based on the idea of the human-being graphical behavior, previously depicted in Subsection 4.2.3 (Figure 7).

Before formalizing the aforementioned property, it is essential to clarify its source, which regards the fact that almost all GCs of the unified human-being model (as identified in Subsection 4.2.3) carry the notion of attainment and impairment of psychological human needs. This notion was linked to other psychological aspects by the proposed graphical behavior, which states that a person's personality defines what human needs he/she chases, and also that the attainment or impairment of these needs entails emotions, and consequently, sentiments (the descriptions of all these psychological aspects were previously presented at the beginning of Chapter 2). Given these findings, the "Aggregated psychological essences" property of the unified human-being model is proposed as follows:

The "Aggregated psychological essences" property

The psychological essences of an individual can be modeled as an aggregation of short, mid, and long-term aspects that are built based on this individual's historical occurrences of attainments and impairments of human needs, where:

- Short-term aspects: regard the attainment or impairment of a human need and its entailed emotions (positives or negatives).
- Mid-term aspects: are the sentiments (positives, neutrals, or negatives) built based on the occurrences of positive or negative emotions segregated by human needs.
- Long-term aspects: accounts for the personality traits that define the human needs chase pattern, encompassing priorities between different needs.

According to this property, if a method could systematically formalize its essential beginning (i.e., the attainment and impairment of human needs), the GCs of the unified human-being model would be more easily identified on usage data by considering the property's findings, where the attainment and impairment of human needs allow the identification of additional short, mid, and long-term psychological aspects. Note that this property's essential nature is fostered by the StPos7, which states, "when a person desires to acquire things or be in situations of satisfaction; the wish to attain human needs". Meaning that the desire to be satisfied according to human needs is an essence, where different humans present different desires (entailed by their distinct personalities). Interestingly, with such a formalization, all GT's of the unified human-being model would be covered, as short, mid, and long-term aspects are approached by the "Aggregated psychological essences" property.

In view of the aforementioned facts that, (1) the unified human-being model presents more descriptions than the unified player model, (2) it is assumed that the unified player model is inherently linked to the unified human-being model, (3) the "Aggregated psychological essences" property allows the exploration of all general human-being model's GTs, and (4) a similar input format between the unified models, the unified human-being model is the adopted one in this thesis to generate the psychological profile of players. This profile encompasses for each player his/her sentiment (positive, neutral, or negative) based on the identified emotions (positive or negative) according to the attainment or not of human needs to a given time-span, together with his/her personality (i.e., the way that this player plays; what he/she seeks and how he/she seeks). The mapping between each psychological aspect and its referenced GCs and GTs is presented in Table 30, totalizing a coverage of 41 GCs and all the 3 GTs of the chosen model. Note that this referencing of 41 GCs regards 40.59% of all 101 GCs, which carries this thesis's notion regarding the best trade of between input simplicity and GCs' coverage, suggesting that increasing coverage would also increase the input complexity together with new identification rules.

The definition of each referenced GC of the proposed profile was also based on the related works presented in Section 2.7. Where the work of (BOSTAN, 2009) presented tips for the first needed step, as it provides procedures to straightforward link actions in-game with the attainment of human needs of the Murray model (the six needs considered in the proposed profile, which regards the notions of Materialism, Power, Affiliation, Achievement, Information, and Sensual). It is important to highlight that this linkage between actions and human needs is the most straightforward linkage between GCs and data identified in this thesis. In addition, the work of (POPESCU; BROEKENS; SOMEREN, 2014) provides a means to identify emotions occurrences regarding eight internal emotions (i.e., hope, fear, joy, distress, satisfaction, fears-confirmed, disappointment, and relief) and eight social emotions (i.e., anger, shame, gratitude, gratification, happy-for, pity, gloating, and resentment) of the OCC model (that are linked to the 15 referenced emotions GCs)[2]. Also,

[2] The GCs linked to each emotion is presented in Subsection 5.3.2.

Table 30 – Mapping between each psychological aspect of the proposed psychological profile and its essential referenced GCs and GTs

Psychological Aspect	**Referenced General Characteristic (GC)**	**Referenced General Topic (GT)**
Human need	PsyHM1, PsyHM4, PsyHM5, PsyHM6, PsyHM7, and PsyHM8	Short-term psychological aspects
Emotions	StNeg3, StNeg5, StNeg7, StNeg8, StNeg9, StNeg10, StNeg11, StPos1, StPos2, StPos3, StPos4, StPos5, StPos6, StPos7, and StPos11	Short-term psychological aspects
Sentiments	MtPos1, MtNeut1, MtNeg1, SelfPer1, SelfPer2, SelfPer3, and SelfPer7	Mid-term psychological aspects
Personality traits	LtDecMak2, LtGen4, LtGen6, LtGen9, LtGen10, LtGen11, LtGen12, LtGen13, LtGen14, LtGen15, LtGen16, LtGen17, and LtGen18	Long-term psychological aspects

procedures to identify sentiments and personality traits are proposed based on literature findings regarding the identified "Aggregated psychological essences" property, entailing in the referencing of the seven GCs of the sentiment aspect and the 13 GCs of the personality trait aspect (the literature support to this property is presented in the Method Assessment Section 5.5).

It is essential to highlight that the structure depicted by Table 30 is our starting point to extract psychological metrics from data. Meaning that the proposed psychological profile of this thesis is compound of two parts, one containing the metrics values, and the other where these metrics are explained, where the GCs linked to each metric represent its descriptions. Bearing in mind the chase of proposing metrics to represent the identified psychological aspects of Table 30, in a situation where these aspects are analyzed individually (i.e., assuming no linkage between them), each aspect could be approached by considering only its referenced GCs. By contrast, if it is assumed a linkage between the psychological aspects, combinations of GCs from different aspects can be performed to generate enhanced metrics, which is the case of the proposed method presented in Section 5.3. Given this, the mapping depicted in Table 30 is the basis of an analysis that proposes enhanced metrics that links notions of different psychological aspects to provide more insightful pieces of information regarding human behavior (such as the metrics depicted in Subsection 5.3.2). Therefore, the resultant set of psychological metrics of the proposed method encompasses linkages between the aspects of human needs, emotions, sentiments, and personality. The next sections give more details about the method implementation that generates the metrics of the proposed profile, highlighting the GCs linked to each one of these metrics.

5.2 Method Assumptions

An essential aspect of the proposed method is that it is built to deal with a massive number of players (as described in the Scope section 1.5), being not an experimental approach applied to a “home-made game” with few players. Because of it, new positive and negative aspects can be added to the RQ9 answer, such as follows:

- Negative aspects
  1. As we work with usage data of released games, we can only use the information that is available to us, being not possible to modify the game to obtain desirable features.
  2. The assessment of psychological aspects in humans is done in two main ways, (1) through the appliance of questionnaires and (2) through the observation of a specialist (LANKVELD, 2013). In our case, due to the nature and volume of the data, both assessment techniques cannot be applied, as it is not possible to fit the usage data to the questionnaires’ answers nor manually evaluate the behavior of a massive number of players with a specialist. Therefore, the psychological aspects identified in this work are theoretical, being not possible to assume them as true (the real state of a player in a given time). However, this theoretical information can identify and improve risk prediction models (as shown next, in the positive aspects).
- Positive aspects
  1. A massive addition of information in the Game Analytics field, allowing the discovery of new players’ trends.
  2. After the first running of the method, it allows the automatic identification of several psychological aspects over time.
  3. An inner description of players’ motivations, allowing to depict what aspects of the game content please the players more, and also which ones displease more, avoiding empirical considerations. This kind of information can be used to plan new game contents, evaluate the acceptance of game upgrades, and improve risk prediction models’ accuracy.
  4. The identification of reasons for abandonment, entailing in labeling possible abandonment profiles to players before they abandon the game or start to show lack of motivation. For example, after a player losing several PvP battles, the impairment of his/her sentiment about fighting (which turned from positive to negative) led him/her to leave the game. Note that this information is useful for countermeasures, as it can be used as an apriori knowledge to identify other

players with similar behavior and what aspect of the game content may lead them to leave the game. Moreover, it can help the development of new game content that avoid or minimize this situation.

5. A detailed aspect of the game content consumption, highlighting when players started to explore the different possibilities offered to them.
6. An overview of the game consumption illustrating the amount of content not yet consumed by players (considering what they like to do). This information helps define the best moment to release a game upgrade and what kind of content could be better.

The proposed method requires usage data with high granularity to identify the psychological aspects, i.e., data that contains players' actions. A candidate game must contain the following characteristics:

- Usage data with:

  Players' ID

  Timestamps

  Players' actions

The following characteristics are desirable but not necessarily required:

- A non-linear story.
- The entertainment as an objective (voluntary usage).
- Usage data containing descriptions of social interaction, such as the IDs of the causer and the affected, and the list of players aware of each interaction.

We consider the non-linearity aspect useful because it allows players to choose inside a range of options, the ones that he/she most like (highlighting his/her psychological aspects). Games too linear (e.g., without optional paths, details in Section 3.3) obligate players to do specific actions to advance in the game plot. In these cases, the proposed method will be applied considering only the psychological aspects presented in the game (e.g., materialism and power related needs) and not the players' choices to do or not them. Only if a player finished the game, it will be possible to assume the player's chase for the offered needs, as he/she theoretically enjoyed the game due to its conclusion (he/she did what he/she wanted; the player's choice).

Besides entertainment games where players are volunteers, we understand that this approach can also be applied to serious or educational games, but with some other conclusions. In these cases, the psychological information identified can be seen as a description of the psychological interactions available in the game content, similar to the situation of games too linear. However, it is impossible to assume that a player chose to play to attain some human needs, even though some of them may be attained during the gameplay, as the main reason to play could be an obligation and not a leisure activity. Therefore, it is impossible to assume the psychological information of these kinds of players as their own essence (their own will).

The information regarding social interactions in usage data is desirable but not required. This kind of information allows identifying social emotions, unlike the internal ones that do not require such aspect. More details about it are described in Section 5.4.

5.3 Method Overview

The method proposition has two central moments. The first where the basic concepts are understood and then a fundamental structure of the method is proposed (Section 5.3.1), and the second where each step of the method is approached (Section 5.3.2). As a summarized view, Section 5.3.3 presents a straightforward running example of the metrics computations given some input data regarding two players.

5.3.1 Method Fundamental Structure

The method starting point is the usage data, which depicts all the players' choices (actions) in-game. We understand that these choices illustrate psychological aspects according to the concepts of the identified "Aggregated psychological essence" property, depicted in Section 5.1.

Inside the game content universe, the actions done by a player are guided by his/her personality traits that define which human needs to seek, whereupon the attainment or not of such needs entails an emotion, which in turn can entail a sentiment (after a regular occurrence; the definition of "regular" is explained further) that affects the players' motivation to continue playing. However, we assume that a bias about this idea may exist, as in some cases, players may not chase their most desired human needs when the game mechanisms and environment are not mastered. Therefore, we assume that at the beginning of a game, players (the new ones) have an exploratory behavior to identify the possible interactions (learning phase), which can be seen as an "unusual" behavior. After some time, as they learned more about the game, they behave driven by their own psychological aspects (doing more what they want). To label players as more or less "experts" about the game mechanisms and environment, the Commitment metric can be used (see Section 3.7 for more details) (KUMMER; NIEVOLA; PARAISO, 2017b; KUMMER; NIEVOLA; PARAISO, 2018b; KUMMER et al., 2016), demanding an extra input regarding the players' score. Additional considerations about the new players' bias are depicted in Section 5.4.1. Figure 26 illustrates the overview of the proposed method.

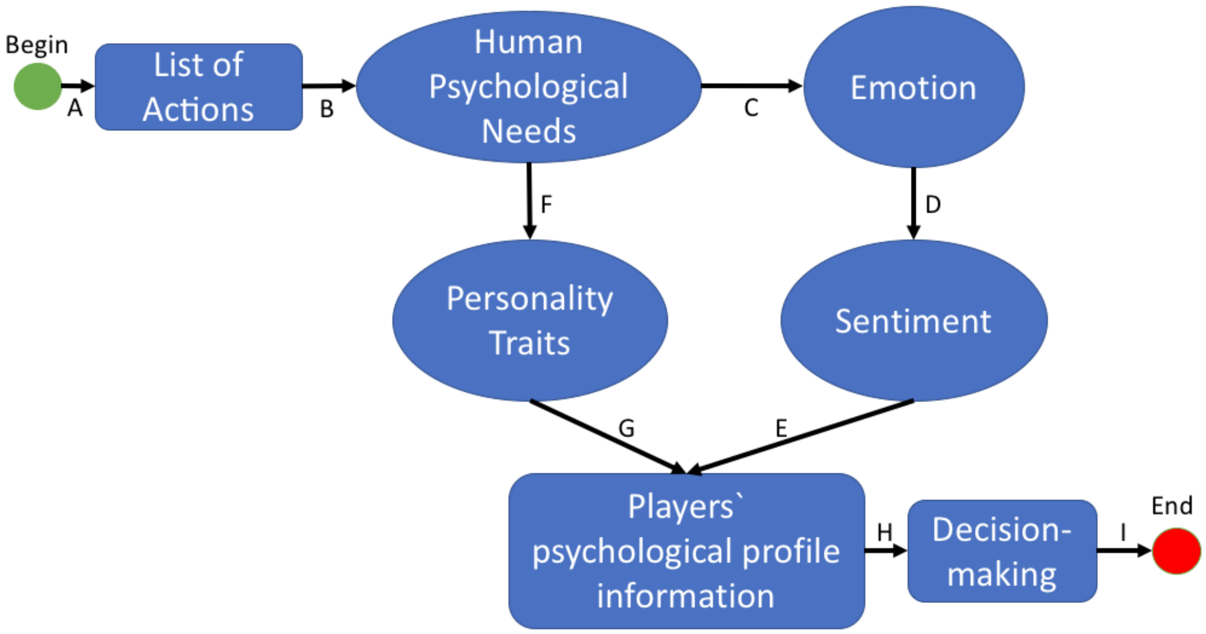


Figure 26 – Method overview

The letters attached to each arrow are labels to identify each connection (or transition); the alphabetical order does not apply here. The rectangular steps regard data preparation or actions done by specialists, while the rounded steps represent the psychological profile identification process.

As we can see, the proposed method generates a psychological profile of players containing only the referenced aspects depicted by Table 30. Next, each step of the proposed method is explained.

5.3.2 Method Steps

Transition A can be understood as the data collection procedure, followed by its preprocessing (for more details, please see Section 3.6.2). Optionally, the Commitment metric can be added in this transition to verify the players' behavior bias (for more details, please see Section 3.7). An example of a list of actions is shown in Table 31.

The transition B is the only manual step of the proposed method. It regards the manual association of each possible action in-game with one or more human needs. We named this association the "action-need map". This map is implemented following the descriptions proposed by (BOSTAN, 2009) and presented in Section 2.7. In addition to it, we assume that an action may have a positive or negative influence, as it can attain or impair human needs. After building this map, its application can identify all players' human needs attained or impaired on each timestamp according to the performed actions.

Table 31 – Example of a list of actions with the optional Commitment metric

Player ID	Timestamp	Player Level	Commitment Degree	Action Done
1	12:32:00 – 01/02/2018	1	Low	Level up
1	12:33:12 – 01/02/2018	2	Low	Get item
2	12:33:37 – 01/02/2018	25	Average	Join guild
3	14:01:03 – 01/02/2018	60	High	Acquire quest
3	14:05:55 – 01/02/2018	60	High	Loot
3	14:08:11 – 01/02/2018	60	High	Complete quest

Note that different games may have distinct action-need maps. The human needs chosen in this method are the ones of the general human-being model that links to the Murray model (MURRAY, 1938) (such as depicted in Table 30). Hereafter, these needs are referenced as Materialism, Power, Affiliation, Achievement, Information, and Sensual to account to their main idea presented in their descriptions, such as respectively depicted by the following GCs: PsyHM4, PsyHM5, PsyHM1, PsyHM6, PsyHM7, and PsyHM8. One example of an action-need map can be seen in Table 32, where positive polarity means attainment and negative polarity impairment.

Table 32 – Example of an action-need map

Action	Human Need	Influence Polarity
Get exp	Achievement	Positive
Get item	Materialism	Positive
Die	Power, Achievement	Negative
Kill	Power	Positive
Join guild	Affiliation	Positive
Leave guild	Affiliation	Negative

The emotions adopted in transition C regards the ones of the unified human-being model that links to the OCC model (ORTONY; CLORE; COLLINS, 1990). It was chosen because these emotions allow a logical simulation, as done by (POPESCU; BROEKENS; SOMEREN, 2014) (presented in Section 2.7). This method similarly uses these logic concepts, but with another objective, the identification of emotions. To allow such an identification per human needs and also encompass social aspects, the referenced emotion GCs (of Table 30) were analyzed together with the ones that reference the six adopted human needs, social interactions, and self-perceptions, being all of them grouped by emotion names, as similarly proposed by (ORTONY; CLORE; COLLINS, 1990). Table 33 shows the group of GCs for each emotion name, dividing them as internal or social emotions, portraying also each emotion polarity (positive or negative). Each emotion description can be seen as the composition of all its GCs descriptions. Note that all internal emotions carry short and long-term notions, as their linked occurrences (short-term) regards what a person wishes (long-term). Moreover, besides these short and long-term aspects, some social emotions also carry notions regarding mid-term aspects, which are linked to liking or not other persons (mid-term) and self-perception of doing something good or not toward others (mid-term). It is important to highlight that, even though the proposed names are similar to the OCC model's ones, their descriptions regard the unified model's associated GCs, which present enhanced details.

Table 33 – Considered emotions

Emotion Name	Emotion Type	Polarity	General Characteristic (GC)
Hope	Internal	Positive	StPos7, StPos11, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Fear		Negative	StNeg5, LtGen9, LtGen10, LtGen11, LtGen13 LtGen14, and LtGen15
Satisfaction		Positive	StPos3, StPos7, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Fears-confirmed		Negative	StNeg9, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Disappointment		Negative	StNeg10, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Relief		Positive	StPos4, StPos7, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Joy		Positive	StPos8, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Distress		Negative	StNeg3, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Anger	Social	Negative	LtDecMak2, StNeg3, LtGen6, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Shame		Negative	LtDecMak2, StNeg11, LtGen4, LtGen13, SelfPer3, and SelfPer7
Gratitude		Positive	LtDecMak2, StPos6, LtGen4, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15
Gratification		Positive	LtDecMak2, StPos5, LtGen4, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, LtGen15, SelfPer1, SelfPer2, and SelfPer7
Happy-for		Positive	LtDecMak2, StPos1, MtPos1, LtGen4, and LtGen13
Pity		Negative	LtDecMak2, StNeg7, MtPos1, LtGen4, and LtGen13
Gloating		Positive	StPos2, MtNeg1, and LtGen10
Resentment		Negative	StNeg8, MtNeg1, and LtGen10

The emotion identification process demands a time-span definition (e.g., daily, weekly, monthly, or per each game session). The time-span size is usually defined by the interested person (e.g., the game producer), which considers a desirable rate of monitoring of players' motivation. Therefore, the time-span size is configurable in the proposed method.

Given the division between social and internal emotions, for each time-span and for each player, the internal emotions are identified based on player's goals, their likelihood, and their final result, while the social emotions are determined by occurrences of events (desirable or not) that attained or not a self's human need or another's one, being they caused by others, or that affected others (more details about the emotions' identification process are presented in Section 5.4). In addition, a player's goal is assumed as a human need that one wants to attain. Concluding, the transition C final result is a list of emotions linked to each human need (an example is shown in Table 34). Note that, depending on game content, a game may not affect some kinds of human needs (e.g., Sensual); different games may affect different needs.

Table 34 – Example of emotions identified in a given time-span

		Power		**Affiliation**	
Player ID	**Time-span**	**Positive**	**Negative**	**Positive**	**Negative**
1	1st time-span	Hope, Satisfaction, Satisfaction, Joy	-	Relief	Fear, Distress
1	2nd time-span	Hope	-	Hope	Fears-confirmed, Fears-confirmed
2	1st time-span	Hope, Joy	-	-	Fear
2	2nd time-span	-	-	-	-

Transition D regards the computation of sentiments for each player according to his/her list of emotions. As far as our knowledge goes, there is no metric to compute a sentiment; therefore, we proposed to compute it into two perspectives, time-span based and historical, where each one has six values, one for each human need (i.e., Materialism, Power, Affiliation, Achievement, Information, and Sensual). It is essential to highlight that the hope and fear emotions are not considered in the sentiment computations as they regard expectancy of players having more successes than failures in a given time-span according to each human need (as shown further in Subsection 5.4.3.1) [3].

[3] This fact can be justified given the following example. Let us assume that a person has hope of winning a battle and loses it. If the hope emotion was considered, the final sentiment would be neutral (due to the balance between one positive (hope) and one negative emotion (disappointment)); however, it should be negative, as the only attempt was not successful. The same rule applies to the opposite, where a person can have a fear of fighting and wins, entailing a neutral sentiment according to the balance between one negative emotion (fear) and one positive emotion (relief); however, it should be positive due to the unique and successful attempt.

A time-span based sentiment of a player is proposed as the sum of all positive emotions divided by all emotions of a given time-span according to a human need (disregarding the hope and fear emotions, as previously mentioned), such as depicted in Equation 5.1.

$$\text{Sentiment}_{\text{time-span}} = \frac{\sum_{i=1}^{n} PositiveEmotion_i}{\sum_{j=1}^{m} Emotion_j} \tag{5.1}$$

Where n is the total number of positive emotions according to a human need, m the total number of emotions according to a human need, $PositiveEmotion_i$ has value 1 and regards a positive emotion occurrence $_i$, and $Emotion_j$ has value 1 and regards an emotion occurrence $_j$. Note that all values regard the same time-span. This sentiment metric has a range between 0 and 1, where a value below 0.5 means a negative sentiment, above 0.5 a positive sentiment, and equal to 0.5 a neutral sentiment.

Moving to the historical sentiment perspective, it is computed as the mean value of all time-span based sentiments according to a human need. It means that, assuming a daily perspective, if a player played for ten days, his/her historical sentiment regards the mean value of all his/her ten daily sentiments. Equation 5.2 formalizes this historical computation.

$$\text{Sentiment}_{\text{historical}} = \frac{\sum_{i=1}^{n} Sentiment_{time-span\_i}}{n} \tag{5.2}$$

Where n is the total number of time-spans played, and $Sentiment_{time-span\_i}$ is the sentiment computed according to the time-span $_i$. This metric has the same range as the previous one, between 0 and 1, where a value below 0.5 means a negative sentiment, above 0.5 a positive sentiment, and equal to 0.5 a neutral sentiment.

It is essential to highlight that the sentiment computation creates for each player and for each time-span a set of 12 sentiment values, such as follows:

- From the time-span perspective:

  Materialism sentiment.

  Power sentiment.

  Affiliation sentiment.

  Achievement sentiment.

  Information sentiment.

  Sensual sentiment.

- From the historical perspective:

  Materialism sentiment.

Power sentiment.

Affiliation sentiment.

Achievement sentiment.

Information sentiment.

Sensual sentiment.

Note that a player can repeat the same action in the same time-span many times, which gives the idea of frequency. Assuming a player killed another one nine times (generating a positive emotion nine times) and got killed by him/her once in their last duel (entailing in a negative emotion), we still assume a positive sentiment due to the value of 0.9 obtained from the $Sentiment_{time-span}$ computation (9 positive emotions/10 generated emotions = 0.9). In conclusion, the proposed Sentiment metrics do not consider the order of occurrences of emotions, as we consider a sentiment toward something as a representation of the sum of emotions' occurrences and not of the last situation. If one wants to check if the last occurrence of an action entailed a positive or negative psychological aspect, only the last emotion's polarity could be used, not needing to compute any Sentiment value.

Sentiments are generated after a regular occurrence toward something (BEN-ZE'EV, 2000; RUSSELL; BARRETT, 1999; MUNEZERO et al., 2014). In our approach, this occurrence is defined by the frequency of positive and negative emotions, which are interpreted in two perspectives, time-span and historical, where each one considers the six types of human needs. Table 35 illustrates the two perspectives values regarding the emotions depicted in Table 34. Note that no sentiment is computed when a player did not play in a given time-span (such as happened to player 2 in the second time-span). It is important to highlight that a neutral sentiment can occur in two circumstances: (1) when there are no emotions regarding a human need and (2) when the number of positive and negative emotions are the same.

Table 35 – Sentiment computed based on Table 34

		Time-span Sentiment		**Historical Sentiment**	
Player ID	**Time-span**	**Power**	**Affiliation**	**Power**	**Affiliation**
1	1st time-span	1 (Positive)	0.5 (Neutral)	1 (Positive)	0.5 (Neutral)
1	2nd time-span	0.5 (Neutral)	0 (Negative)	0.75 (Positive)	0.25 (Negative)
2	1st time-span	1 (Positive)	0.5 (Neutral)	1 (Positive)	0.5 (Neutral)
2	2nd time-span	-	-	-	-

As a final remark about sentiments, these metrics computations were based on the following GCs of the unified human-being model: MtPos1, MtNeg1, MtNeut1, LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15.

The identification of personality traits in the transition F is based on the concept that personality is an almost immutable structure of the human-being that drives our choices (LANKVELD, 2013; CARVER; SCHEIER, 2012). Thus, we assume that if we look at the choices made by a player (assuming choices as the attempts to attain human needs; transition B idea), we can identify patterns of behavior, conjecturing them as personality traits. Two perspectives of analysis are proposed to identify such patterns, one where the chase priority of human needs is approached (named as Macro Spectrum), and the other where the sequence of players' actions are accounted for (the called Micro Spectrum), where each perspective has its set of metrics (detailed in Section 5.4). These perspectives are approached to cover the gap present on the non-invasive approaches, where it is not possible to ask players to fulfill a questionnaire or use the knowledge of a specialist to identify inside a massive amount of players their traits. Regarding the Micro Spectrum, it is based on the common "Game Paths" followed by players. Both Micro and Macro Spectrums' metrics are considered as personality traits. Moreover, we understand that the longer a behavior is observed, the more described and identifiable it is; thus all the time-spans are analyzed together in this step. As an additional remark, the linked GCs to the Macro and Micro Spectrums' metrics are as follows: LtGen9, LtGen10, LtGen11, LtGen13, LtGen14, and LtGen15.

Transitions E and G aim to group the psychological metrics obtained from the sentiment and personality trait steps. Note that the metrics regarding human needs and emotions are not provided to the decision-makers as they are used to compute the sentiment and personality trait ones, which, inherently, carry their essences. From this moment, the **players' psychological profile** is finished, providing to the decision-maker psychological information regarding the players' sentiments and personality traits. Indicating individually for each player and for each time-span, what aspects of the game content players prefer regarding their human needs chase patterns (i.e., personality traits), as well as their feelings (i.e., sentiments) in consuming such desirable contents. In addition, Table 36 summarizes the number of metrics per psychological aspect, even though some of them are only approached in the further Subsection 5.4.

Besides the set of metrics present in the proposed psychological profile, new metrics can be proposed based on them. A general overview of players' historical sentiments according to the game content consumption can be given by the sum of all players on each polarity of each human need. To summarize this computation, the general Equations 5.3, 5.4, and 5.5 are proposed, where n is the total number of players, $PlayerPositive_i$, $PlayerNeutral_i$, and $PlayerNegative_i$ correspond to the sentiment of player $_i$ and have

Table 36 – Number of metrics computed per psychological aspect

Psychological Aspect	**Number of metrics**
Human needs	12 metrics, where: - Six regard the attainment and six the impairment of each considered human need (i.e., Materialism, Power, Affiliation, Achievement, Information, and Sensual).
Emotions	71 metrics divided into two groups: - Internal emotions (48 metrics), where: - Six groups regarding each human need containing each eight metrics referencing the following emotions: Hope, Fear, Joy, Distress, Satisfaction, Disappointment, Relief, and Fears-confirmed. - Social emotions (23 metrics) divided into three groups, where: - The first one contains 18 metrics referencing the Anger, Gratitude, and Gratification emotions in relation to all the six human needs. - The second contains 3 metrics referencing the Shame, Happy-for, and Pity emotions in relation to the Affiliation need. - The third contains 2 metrics referencing the Gloating, and Resentment emotions in relation to the Power need.
Sentiments	12 metrics, where: - For each one of the six human needs, there are two metrics regarding the time-span based and historical sentiments.
Personality traits	48 metrics, where: - One regards the current Game Path. - 36 regard the human needs priorities chase, where: - For each one of the six human needs, there are six possible positions in a priority ranking. Each position (i.e., first, second, third, fourth, fifth, or sixth) contains the percentage of time where a given human need was placed into this position. - Six are boolean values indicating if each one of the human needs were chased or not. - Two encompass similarities (from the Macro and Micro Spectrums). - One regards a personality influence over the sequence of actions. - Two consider the amount of psychological available content.

value 1 when it is positive, neutral or negative respectively, otherwise the value 0 is assumed. Furthermore, these equations are applied for each human need. Figure 27 shows an example of this overview.

$$\text{GeneralPositiveSentiment} = \sum_{i=1}^{n} PlayerPositive_i \tag{5.3}$$

$$\text{GeneralNeutralSentiment} = \sum_{i=1}^{n} PlayerNeutral_i \tag{5.4}$$

$$\text{GeneralNegativeSentiment} = \sum_{i=1}^{n} PlayerNegative_i \tag{5.5}$$

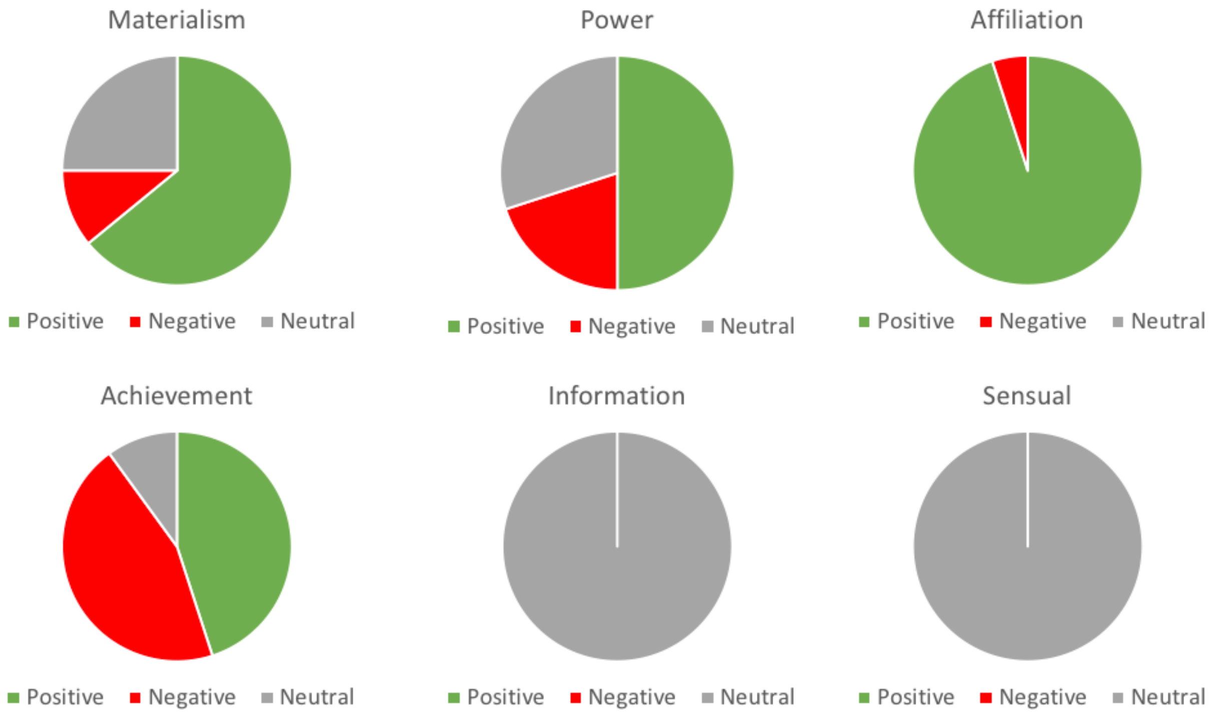


Figure 27 – Example of the sentiment general overview for a given time-span

According to the hypothetical situation depicted in Figure 27, the game content pleases more than half of the active players regarding the Materialism need, pleases half of them in the Power need, has a good perspective in the Affiliation need (almost 100% of positive sentiment), presents a balanced perspective in the Achievement need although it has the higher negative aspect over all human needs, and does not affect the Information and Sensual needs of players. This overview can be seen as a "snapshot" of the players' sentiments since their first time-span. It is essential to highlight that pleasing and entertaining may have different meanings in this context. For example, games too easy, where players attains regularly their needs (i.e., regularly pleased; "greener snapshots") can be less entertaining compared to games that are more difficult, which present more negative sentiments ("redder snapshots"). Therefore, what is entertaining or not is linked to the expected difficulty degree of the active players in attaining their needs; however, it is expected that, when a game has its difficulty degree in a desirable level according to its active players, the "greener" a game snapshot is (even in very difficult games), the more entertained its players are.

To allow the chase of the ideal scenario of "greener" snapshots carrying the notion of entertainment, it is first needed to identify the ideal challenge degree of a game. To do such a identification, the sentiments of players can be associated with the churn rate in view of game upgrades. For example, if the churn rate increased together with the positive sentiments increase, it means that the players did not approve the new challenge (too easy), being preferred a more difficult one; by contrast, if the negative sentiments increased, it means that the players did not approve the new challenge because it is too difficult, being preferred a more easier one. Differently, if the churn rate decreased together with the positive sentiments increase, it means that players are enjoying and accomplishing the new challenges, meaning that the expected challenge degree was achieved; by contrast, if the negative sentiments increased, the players are enjoying the new difficult challenges, meaning that the expected challenge degree was also achieved. Given this example, it is possible to see that looking at the historical behavior of the churn rate and the sentiments of a game, a preferred challenge degree can be identified. Note that if only the churn rate was considered in this analysis, its variances would not give sufficient information about the ideal challenge degree, as it lacks the notions of how players are interacting with the game content (i.e., their sentiments). It means that to turn a challenge degree nearer to an ideal level, the players' approval in continuing playing (i.e., the churn rate) must be linked to their easiness in attaining their needs (i.e., the sentiments), allowing changes in the difficult degree to turn it easier or harder when the no permanence of player in the game is linked to positive (harder is preferred) or negative sentiments (easier is preferred). When the churn rate is decreasing, it is possible to assume that the challenge degree is acceptable, not needing to account for the sentiments. Given that the preferred challenge degree was identified, the chase in turning the players' sentiments "greener", which means attaining the players' needs in a way that they feel competent (the Flow idea) (DECI; RYAN, 1985; DECI; RYAN, 1995; HUIZINGA, 2014), can be gauged.

With this information in hands, a game producer can make-decisions (transition H) due to a better understanding about how his/her game entertains or not the active players, avoiding empirical considerations (SHESTOV, 2018). For instance, assuming that the ideal difficulty degree of a game was already identified, the snapshot of Figure 27 can be used to plan a next upgrade focusing on reducing the negative sentiments (e.g., the Achievement one) or improving the positive ones. The new upgrade can also approach human needs that have not been attained yet, such as the Information (adding a story) and Sensual (adding new kinds of relationship between players, e.g., the possibility to get married). Additionally, this snapshot can assess a game upgrade by verifying the players' new sentiments according to the new content. In conclusion, when the difficulty degree of a game is set to the expectancy of its players, high levels of negative sentiments highlight a disapproval over new contents, while high levels of positive sentiments accentuate the players' acceptance. For a more specific analysis, if desired, only the last time-span sentiments can be considered.

During the game content consumption, each player has access to different possibilities of interaction (the Game Path idea), and consequently, these new opportunities can change the current sentiment of players about a human need. The game consumption idea is that, sooner or later, each part of the game content will be reached by players, and then it can entertain them (to generate a positive sentiment linked to an expected challenge degree) or not. Moreover, nothing prevents a player with a history of positive sentiments to change his/her perception of the game when faced with negative emotions. This fact highlights the importance of the continuous monitoring of players' behavior.

Finally, the last transition (I) regards the appliance of the decisions made and their monitoring. The overview of the proposed method, depicted in Figure 26, contemplates steps done before and after the generation of the psychological profile of players (the rounded steps); however, it is essential to highlight that this work is focused on obtaining the proposed psychological profile from data and not on solving problems related to data acquisition or the best decision to make; these aspects are left to specialists (as described in the Scope Section 1.5). As a final remark, the proposed method can be reapplied to each new time-span (a new iteration), allowing in that way the portraying of the historical overview of the players' sentiments and personalities.

5.3.3 Straightforward Running Example of the Psychological Metrics Computations

In this running example, we consider two hypothetical players named Johann and Maija during three consecutive time-spans, Day 1, Day 2, and Day 3. The hypothetical game has the following five possible actions to be performed by players: Buy Item, Kill Another Player, Die, Collect Ore, and Lost Item. Bearing in mind all the transitions presented in Figure 26 (A, B, C, D, E, F, G, H, and I), each transition data will be described. This is an overview that does not depict the internal computations' details but the generated values on each transition. More details about these internal computations are presented in the Sections 5.4.2, and 5.4.3.

Transition A regards the data collection, which is depicted by Table 37. To produce the transition B data, an action-need map is required. Table 38 presents the considered game's hypothetical action-need map (note that this game only offers content regarding two human needs, Materialism, and Power). By applying this map to the input actions, we have the human needs attained and impaired for each player on each time-span, such as depicted by Table 39 (the transition B data).

Table 37 – Running example - Transition A - Players' actions

Player_ID	**Time-span**	**Action**
Johann	Day 1	Buy Item
Johann	Day 2	Die
Johann	Day 3	Kill Another Player
Maija	Day 1	Buy Item
Maija	Day 2	Collect Ore
Maija	Day 3	Lost Item

Table 38 – Running example - Transition A - Action-need map

Action	**Human Need Attained**	**Human Need Impaired**
Buy Item	Materialism	
Kill Another Player	Power	
Die		Power
Collect Ore	Materialism	
Lost Item		Materialism

Table 39 – Running example - Transition B - Human needs attained and impaired

Player_ID	**Time-span**	**Action**	**Human Need Attained**	**Human Need Impaired**
Johann	Day 1	Buy Item	Materialism	
Johann	Day 2	Die		Power
Johann	Day 3	Kill Another Player	Power	
Maija	Day 1	Buy Item	Materialism	
Maija	Day 2	Collect Ore	Materialism	
Maija	Day 3	Lost Item		Materialism

Based on the transition B data, it is possible to simulate the emotions occurrences, which are depicted in Table 40 (the transition C data). Bearing in mind that each emotion carries a polarity, as being positive or negative, the sentiment values can be identified, such as portrayed by Table 41 referencing the transition D.

Table 40 – Running example - Transition C - Emotions occurrences

Player_ID	**Time-span**	**Materialism Emotions**	**Power Emotions**
Johann	Day 1	Hope and Satisfaction	Hope
Johann	Day 2	Hope	Hope and Disappointment
Johann	Day 3	Hope	Fear and Relief
Maija	Day 1	Hope and Satisfaction	Hope
Maija	Day 2	Hope and Satisfaction	Hope
Maija	Day 3	Hope and Disappointment	Hope

The transition F is a little bit more complex as there are different perspectives to measure the players' personality traits. Table 42 shows the Macro Spectrum regarding the

Table 41 – Running example - Transition D - Sentiments occurrences

Player_ID	**Time-span**	**Materialism Daily Sentiment**	**Materialism Historical Sentiment**	**Power Daily Sentiment**	**Power Historical Sentiment**
Johann	Day 1	1	1	0.5	0.5
Johann	Day 2	-	1	0	0.25
Johann	Day 3	-	1	1	0.5
Maija	Day 1	1	1	0.5	0.5
Maija	Day 2	1	1	-	0.5
Maija	Day 3	0	0.66	-	0.5

ranking of needs, whereas Table 43 the chase or not of the two considered human needs, Materialism and Power. Regarding the ranking of needs, when two or more needs were chased in the same degree, they are untied based on alphabetical order. Moving to the Micro Spectrum, it regards the Game Paths identification (also regarding the transition F), which are depicted for each time-span by Figures 28, 29, and 30 regarding Day 1, Day 2, and Day 3, respectively. Table 44 summarizes the players' placement on the Game Path Segments overtime. The similarities between the players' behavior considering the Macro and Micro Spectrums are also computed in transition F, as well as the amount of available content to be consumed by players. Table 45 present the similarity and available content related metrics values.

Table 42 – Running example - Transition F - Macro Spectrum - Ranking of needs

Player_ID	**Time-span**	**1st chased need**	**2nd chased need**
Johann	Day 1	Materialism	Power
Johann	Day 2	Materialism	Power
Johann	Day 3	Power	Materialism
Maija	Day 1	Materialism	Power
Maija	Day 2	Materialism	Power
Maija	Day 3	Materialism	Power

Table 43 – Running example - Transition F - Macro Spectrum - Chase or not of needs

Player_ID	**Time-span**	**Materialism chased**	**Power chased**
Johann	Day 1	True	False
Johann	Day 2	True	True
Johann	Day 3	True	True
Maija	Day 1	True	False
Maija	Day 2	True	False
Maija	Day 3	True	False

The transitions G and E regard the joining of all the psychological metrics regarding sentiments, and personality traits (it encompasses the similarity and content consumption metrics) that we have presented so far in this Section to compose the final psychological

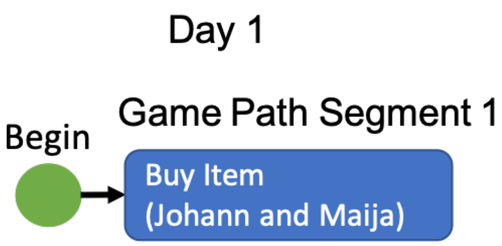


Figure 28 – Running example - Transition F - Micro Spectrum - Day 1

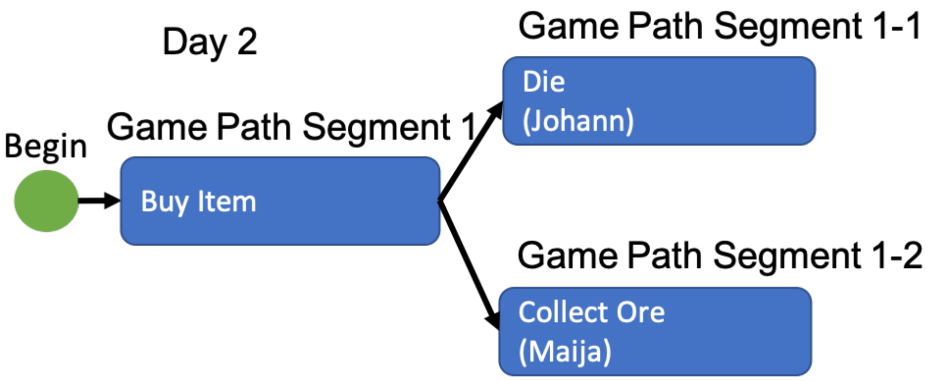


Figure 29 – Running example - Transition F - Micro Spectrum - Day 2

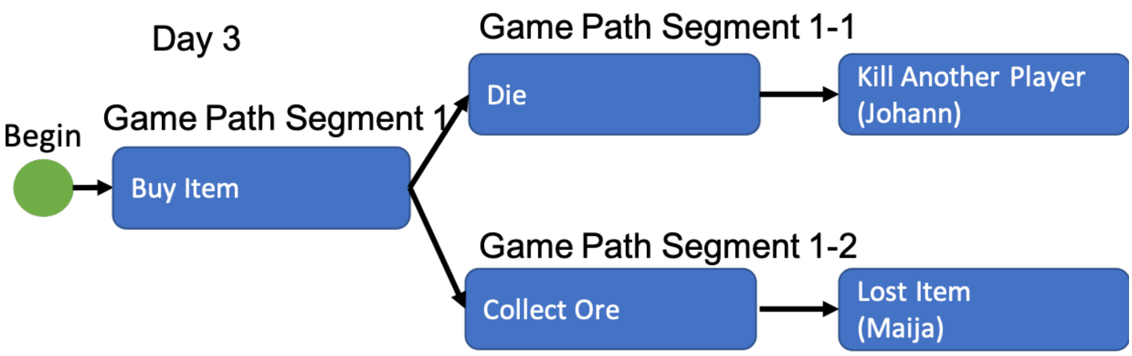


Figure 30 – Running example - Transition F - Micro Spectrum - Day 3

profile of players. This profile can seen then as a combinations of the Tables 41, 42, 43, 44, and 45. The transition H is just a formality to highlight that decisions can be made based on the proposed profile, while the transition I regards the continuous monitoring of the players' behaviors, where more days will be considered overtime, meaning that the method will be re-executed for each new day.

5.4 Special Considerations

This section describes detailed aspects not approached in the method overview (Section 5.3) regarding the bias related to new players (Section 5.4.1), the identification of personality traits (Section 5.4.2), and the identification of emotions (Section 5.4.3).

Table 44 – Running example - Transition F - Micro Spectrum - Players' placement on the Game Path Segments

Player_ID	Time-span	Game Path Segment Placement
Johann	Day 1	1
Johann	Day 2	1-1
Johann	Day 3	1-1
Maija	Day 1	1
Maija	Day 2	1-2
Maija	Day 3	1-2

Table 45 – Running example - Transition F - Similarities in the Macro and Micro Spectrums and the Available Content

Player_ID	Time-span	Macro Spectrum Similarity	Micro Spectrum Similarity	Available Content
Johann	Day 1	100%	100%	0
Johann	Day 2	100%	50%	0
Johann	Day 3	83.25%	50%	0
Maija	Day 1	100%	100%	0
Maija	Day 2	100%	50%	0
Maija	Day 3	83.25%	50%	0

5.4.1 The New Players' Possible Bias

The proposed method has a possible bias associated with the learning process of new players during their "first steps" in a game, as they may not chase their desirable human needs due to a poor understanding of the available interactions inside the game content. Thus, we assume that new players may behave "strangely" compared to older players (i.e., more mature players), who already learned about the game mechanisms and have personal objectives to reach (i.e., human needs). To identify players as less or more mature, the Commitment metric is used, where low committed players are assumed as having the least maturity degree, average committed ones as more mature than the low committed ones, and the high committed ones as having the highest maturity.

We understand that the bias information is relevant to the decision-maker; therefore, we propose to compute a metric to represent it, which we named "Maturity". We assume that the more mature the active players are, the lower the possible bias is. Equation 5.6 depicts the proposed metric.

$$\text{Maturity} = \frac{100 - \frac{P_{avg}}{2} - P_{low}}{100} \tag{5.6}$$

Where P_{avg} and P_{low} represent the percentage of the total active players with an average and a low commitment degree, respectively. The $Maturity$ range is from 0 to 1, where 0 means a low maturity (possible high bias) and 1 a high maturity (possible low bias).

The metric's conception is based on the following idea. Assuming three basic cases where: (1) a $Maturity$ value of 1 happens when a game has 100% of high committed players, (2) a $Maturity$ value of 0.5 happens when 100% of the active players has an average commitment degree, and (3) a $Maturity$ value of 0 is obtained when all active players have a low commitment degree. It is assumed that each commitment degree percentage "pulls" the Maturity metric to their specific zones. Considering this idea, the proposed metric starts assuming all active players as having a high commitment degree to the game (a $Maturity$ value of 1), and then the percentage of average and low committed players are added to this perspective, adjusting the final $Maturity$ value according to their influence (percentage). Looking from the high commitment point of view, as the average committed players "pulls" the metric value to 0.5 and the low committed ones to 0, it means that the "force" of the average committed percentage has half intensity as compared to the low one (this is the reason why the P_{avg} value is divided by 2). Table 46 shows examples of the Maturity metric according to different percentages of players on each commitment degree. Note that values of 0.5 can also be obtained when both percentages of low and high committed players are the same.

Table 46 – Examples of $Maturity$ values according to different percentages of Commitment

Low%	**Average%**	**High%**	$Maturity$
0	40	60	0.8
40	0	60	0.6
60	40	0	0.2
60	0	40	0.4
25	50	25	0.5
20	20	60	0.7
20	60	20	0.5
60	20	20	0.3

5.4.2 Personality Traits Identification Process

We understand personality traits as an almost immutable structure that drives human's choices (long-term aspect). In addition to it, those choices are represented by the players' actions in the game perspective, where each action can attain or impair human needs. Therefore, to identify personality traits in games, we propose to verify, in two different perspectives, how players play according to their choices in-game, where one focuses on the sequences of players' actions (i.e., the generated Game Paths), the called Micro Spectrum, and the other on the human needs chase priorities, the called Macro Spectrum. The next Subsections 5.4.2.1 and 5.4.2.3 approach the Micro and Macro Spectrums, respectively. Also, similarities between players' behaviors on both Micro and

Macro Spectrums can be computed, such as shown by Subsections 5.4.2.2, and 5.4.2.4, respectively.

5.4.2.1 Micro Spectrum - The Game Paths Generation

Depending on the non-linearity degree of a game content, the number of possible Game Paths can vary. However, some Game Paths may not be "truthful" concerning the chase of desirable needs (an influence of the players' personalities) due to the possible bias of new players (Section 5.4.1). Although this bias may exist, we opted to compute the Game Paths regardless of the commitment degree of players because even though the low committed ones may behave "differently" from the others due to their learning process, the way they learn can also be assumed as an influence of their personalities[4]. Another aspect about the players' choices is that they might not have an end, given the continuous nurturing of new game content or the existence of "loops" inside the game (e.g., daily quests). Therefore, we propose to update the identified Game Paths for each new time-span in cases where loops exist or after the release of an upgrade. However, note that while players have not explored all the game content available to them or when there are loops, the Game Paths' update must happen for each new time-span to capture their consumption behavior. In the case where loops do not exist, it is expected that in some moment, after the game content is well known by the players, the identified Game Paths do not change anymore, until the release of a new content.

Besides identifying Game Paths, the placement of players into them is done for every new time-span. It is possible that inside of the same Game Path, different players are in different "steps". Figure 31 illustrates this aspect, where there are three players in the first step, 12 in the second, and only one player at the n step. The player in the n position is called the "head" of the path (it is possible to have more than one player in this position if they behave identically). Note that the concept of the Game Path's end is not used, as Game Paths may grow over time. Instead of it, the "head" concept is adopted to represent the last identified step of a given path. The player in the head position can be understood as a pathfinder, which discovers new possibilities inside the game content which may be followed by other players (with or without their acknowledgment). Moreover, the number of players on each step of a path can portray how much content is yet available to be consumed. To represent this idea, two metrics are proposed, where one measures the amount of available content for a given Game Path (Equation 5.7) and the other the mean value of it per each player of the path (Equation 5.8).

[4] Note that when a new game content is released, all players may experience a learning process while they are discovering the new opportunities, regardless of their previous knowledge or commitment degree.

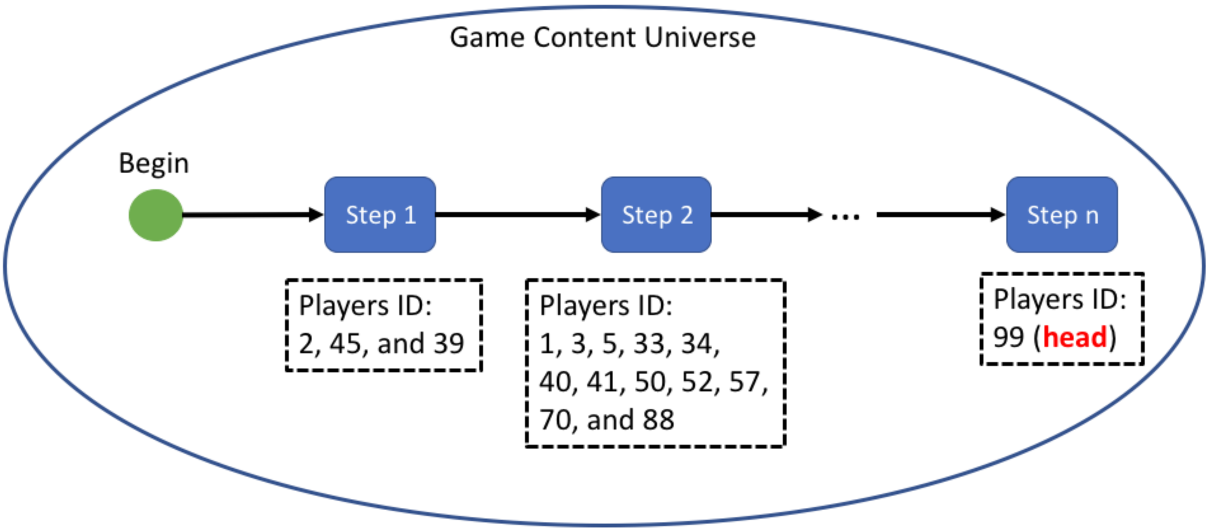


Figure 31 – Example of players in different steps of a given Game Path

$$\text{AvailableContent} = \sum_{i=1}^{n}(\frac{Step_n - Step_i}{Step_n} * PlayersStep_i) \tag{5.7}$$

$$\text{MeanAvailableContent} = (\frac{AvailableContent}{\sum_{i=1}^{n} PlayersStep_i}) * 100 \tag{5.8}$$

Where $Step_n$ is the last step of a given Game Path (the head's step), $Step_i$ is the step in the $_i$ position, and $PlayersStep_i$ is the number of players on the $Step_i$. Basically, Equation 5.7 sums the normalized distance of each player until the head's step[5], while Equation 5.8 calculates the percentage mean value of it per player. In addition to it, these metrics assume that all players inside a Game Path already consumed a part of it, even though a player is in the first step, because to be there, it was needed to consume some content, therefore, the $MeanAvailableContent$ value will never have a value of 100% and the $AvailableContent$ value will vary from zero to the positive infinite. Moreover, every path will have at least one player on it, so it is always possible to compute the metrics. An interesting fact about the $AvailableContent$ metric is that it may be used to identify when a game should have an upgrade due to a shortage of content to be consumed (an alert); thus, the nearer this metric's value is from zero, the more needed an upgrade is. Table 47 presents some examples of the proposed metrics, where it is possible to see changes in their values according to the number of players far from the head's step (to make it clearer, some abbreviations were used where "distance to head" is assumed as DistH and "number of players" as NumP). As a final remark, these metrics only computes the available content regarding the game content already discovered by players. It is essential to highlight that

[5] Assuming, as a result, an AvailableContent of 2, it means that the sum of available content to be consumed in the same manner that other players did (considering the players that are not in the head's position) is equal to 2 * the Game Path length. In other words, the sum of available content corresponds to 2 full passing through the considered Game Path (from its first step until its last one).

these values may change as players discover new interactions in-game. Therefore, the ideal case would be when all possible Game Paths were already discovered, entailing a value without bias.

Table 47 – Examples of the *AvailableContent* and *MeanAvailableContent* metrics for Game Paths with one, two, and three steps

	Step 1		**Step 2**		**Step 3**			
Total of Steps	**DistH**	**NumP**	**DistH**	**NumP**	**DistH**	**NumP**	*Available Content*	*Mean Available Content*
1	0	1	-	-	-	-	0	0%
1	0	100	-	-	-	-	0	0%
2	1	10	0	10	-	-	5	25%
2	1	100	0	10	-	-	50	45%
3	2	10	1	10	0	10	10	33%
3	2	100	1	10	0	10	70	58%
3	2	10	1	100	0	10	40	33%
3	2	10	1	10	0	100	10	0.08%

After the identification of a player's path (i.e., his/her sequence of choices), if it is contained inside of another previously identified path, then this player's path is assumed as being the other one (the bigger), placing he/she in a given step of it, however, if the identified path contains another previously identified path, then, the other path is expanded with the new steps, and the owner of the new steps becomes the new head of it. In addition to it, each Game Path has an identification, which is always kept regardless of expansions or not of it (a special situation regarding the splitting cases is depicted further). These situations are illustrated in Figures 32 and 33. Note that there are two kinds of Game Paths, one regarding the common Game Paths that have an ID and different players placed into them, and another regarding the individual paths (the sequences of choices of each player). In the example presented in Figure 32, its "a" part portraits a common Game Path (where player 85 is the head; the Game Path 1), the "b" part the player 44 individual path, and the "c" part the final arrangement of the Game Path 1, where player 44 was placed in a step of it. For Figure 33, the "a" part also portrays the same Game Path 1, the "b" part the player 44 path (bigger than the previous example), and the "c" part the final arrangement of Game Path 1, where the player 44 becomes the new head of it.

The simplest case is when an individual Game Path of a player is not similar to any other; thus, this path can be assumed as a common one without additional deliberations. Another case is when a common Game Path is splitted, giving the idea of Game Path Segments. A Game Path Segment is a sequence of players' actions that can connect to none or many other Game Path Segments [6]. Hereafter, to encompass the increased complexity

[6] Note that when a Game Path has never been splitted, it contains only one Segment.

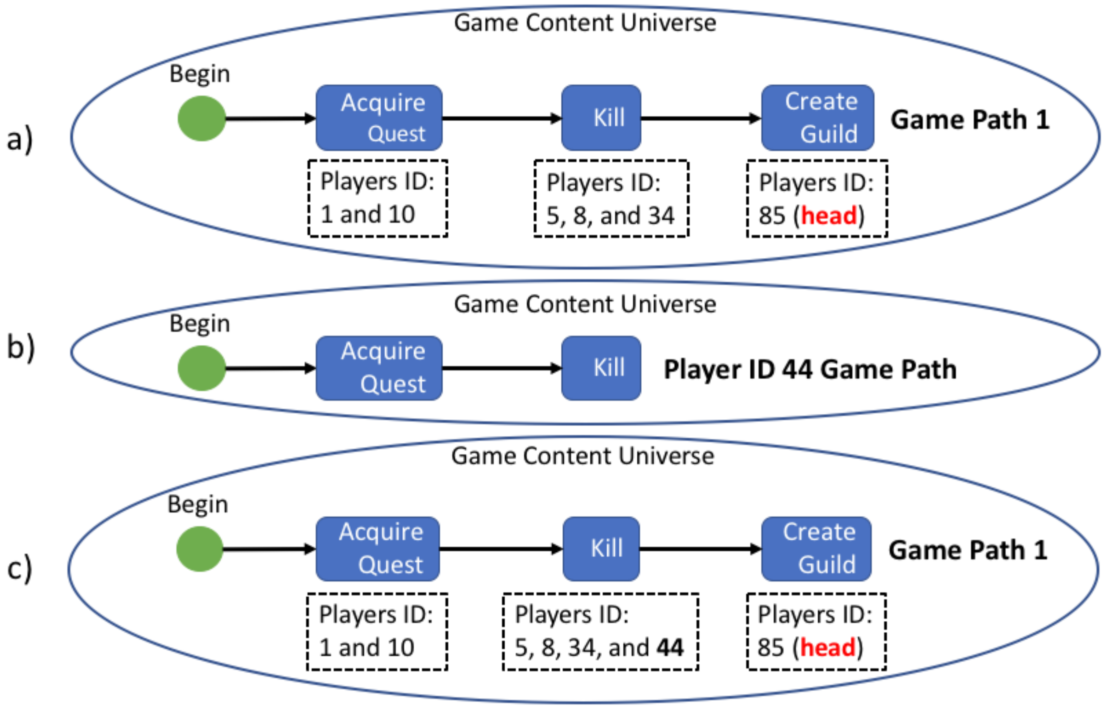


Figure 32 – Example of when a new path is contained inside another one

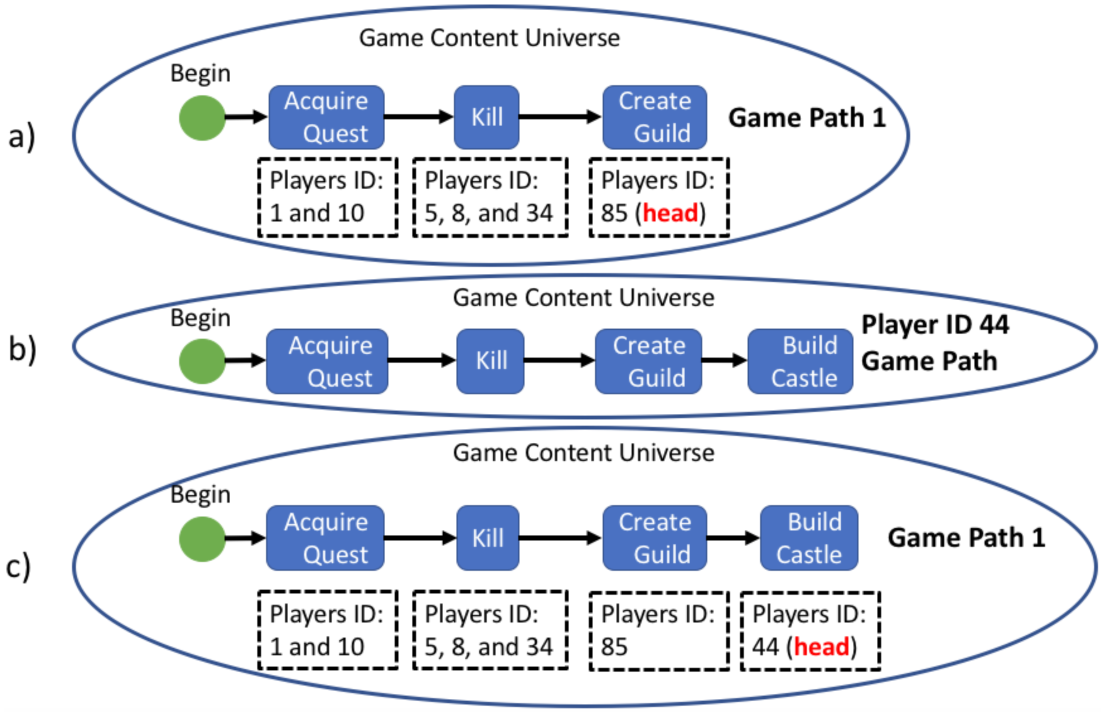


Figure 33 – Example of when a new path expands another one

entailed by the possibility of splitting common Game Paths, the term "Game Path" is then assumed as the structure that contains one or many Game Paths Segments. Note that this term's description is still valid to the previous analysis of this Subsection, and also, that the players' individual Game Paths cannot be splitted, as each one regards a unique behavior. Let us assume the example illustrated in Figure 34, where the "a" part shows a previously identified Game Path Segment 1 (different from other examples), the "b" and "c" parts the Game Paths of players 1 and 2 respectively, and the "d" part the two generated Game Paths Segments (variants) based on the Game Path Segment 1. As in this case, it is not possible to expand the original Game Path Segment, a hierarchy is created, where three different Game Path Segments are present in the same structure[7] (the original one and the two new variants). The original Game Path Segment is maintained without changes, whereas each variant is attached to its end, carrying on their identifications the name of their predecessor (1) followed by an "-" and then a self ID (1-1 and 1-2; note that this approach keeps the original ID). If a splitted segment is splitted again, the same rule is applied, having, as a result, an ID of "1-1-1" and "1-1-2" for example. Note that a split may also happen in the middle of a Game Path Segment, thus in this case, one of the variants will assume part of the steps from the original segment, entailing in a renaming of all possible posterior paths linked to it. According to this representation, it is possible to track the changes of Game Paths hierarchically, highlighting aspects of the game content consumption, such as when there is a game upgrade, and players start to explore the new possibilities (entailing in more splits) or when they avoid doing so (a possible displeasure about the new content).

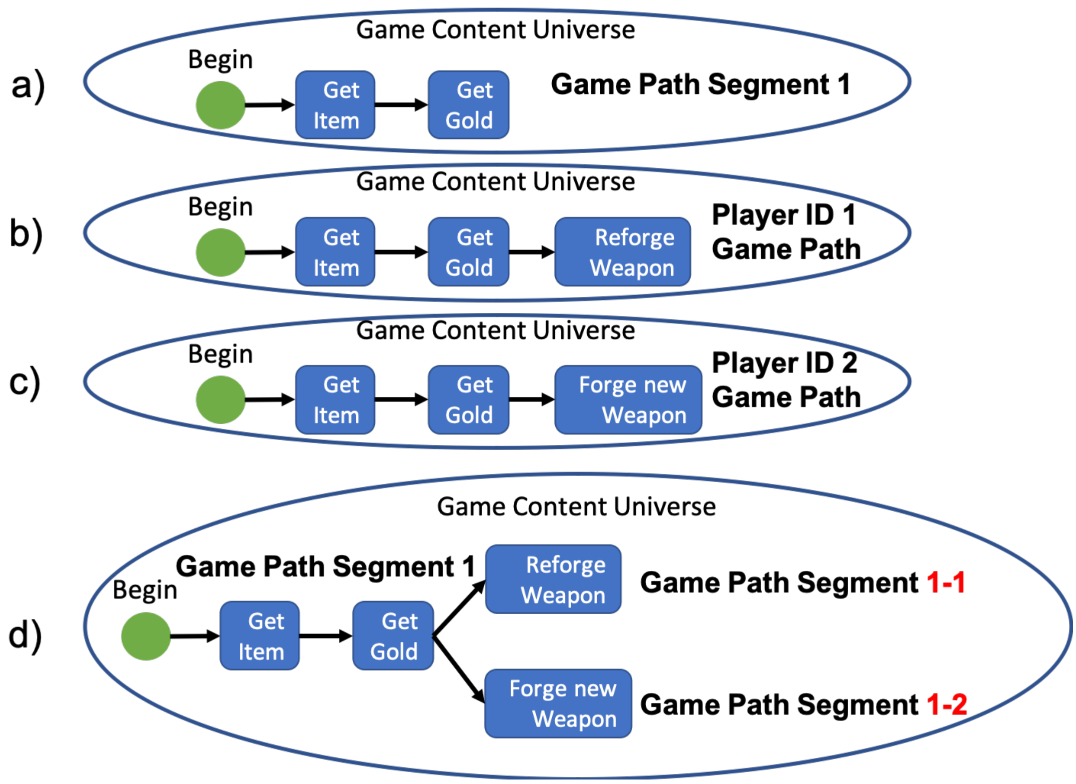


Figure 34 – Example of when a Game Path Segment is splitted

[7] This structure can be referenced as the name of its first Game Path Segment. For example, the Game Path 1 that contains three Segments (Game Path Segment 1, Game Path Segment 1-1, and Game Path Segment 1-2). For the cases where the first segment of a path is being referenced, it is written "The Game Path **Segment** 1".

In the Game Path idea, a player is at only one segment at a time. Thus when a segment is splitted, the players are not duplicated on each new segment. Instead of it, they are kept on their respective segments according to the hierarchy. The hierarchical division of Game Path Segments can also be represented such as depicted in Figure 35, where the steps of each segment are suppressed, showing only the players inside of each one. Moreover, we assume that a segment can have three states: "the root", "original" or/and "variant". An original segment is the one placed before a split, while a variant segment is the resultant segment of a split. Note that for every variant, there is the concept of its original, even though the original is a previous variant. Therefore, inside of a hierarchy of Game Path Segments, there is only one original that is not a variant, the root.

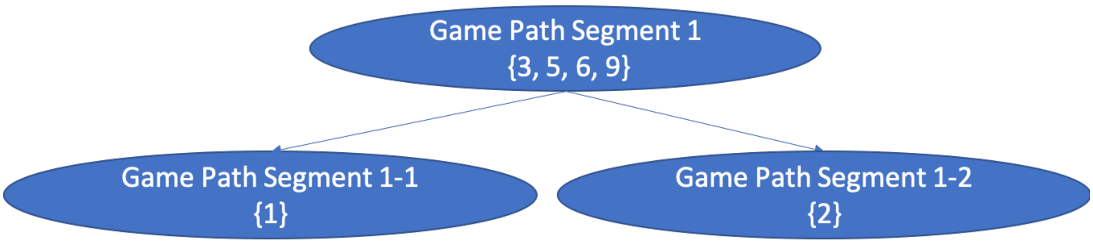


Figure 35 – Example of a Game Path Segments hierarchical division summarized

In view of the possibility to split a segment, the metrics proposed in Equations 5.7 and 5.8 must be revisited to contemplate it. The problem is, it is uncertain which variant a player will follow. Thus, to better highlight the aspects linked to it, let us assume the example depicted in Figure 36.

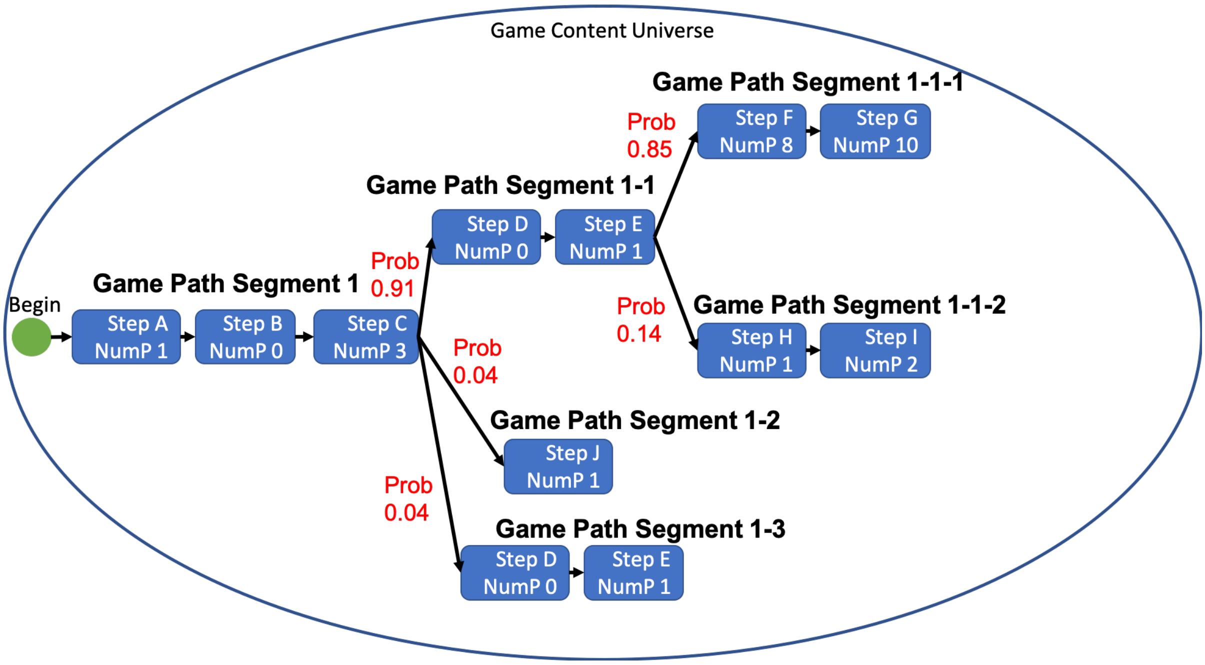


Figure 36 – Example of a Game Path containing six segments with probabilities

Where NumP is the number of players on each step and Prob is the probability of a player following a variant (which is given by the percentage of players that followed it). As we can see, there are four possible full-paths that a player can follow, being they from 1 to 1-1-1, from 1 to 1-1-2, from 1 to 1-2, and from 1 to 1-3. Moreover, depending on the full-path that a player follows, it can have more or less content to consume. Note that the players' choices to follow one or another full-path are not restricted only to the root players. For each possible full-path, each Game Path Segment contained on it can be seen as a part of it. Furthermore, using the example of Figure 36, it is possible to identify three parts for the possible full-path from 1 to 1-1-1, being they 1, 1-1, and 1-1-1. In conclusion, to compute the available content for a splitted common Game Path, it is needed to take into account where the players are (what part of the possible full-path; the segment), and their probability to follow a variant together with the amount of content available on it. The available content of a possible full-path is given by Equation 5.9, and the revisited metrics, *AvailableContentSplitted* and *MeanAvailableContentSplitted*, are presented in Equations 5.10 and 5.11, respectively[8]. As an additional description, Table 48 shows an example of the *AvailableContentSplitted* computation for the example depicted in Figure 36.

$$\text{AvailableContentPossibleFullPath} = \sum_{i=1}^{n}(AC_i * (\prod_{j=1}^{m} Prob_j)) \tag{5.9}$$

Where AC_i is the available content (Equation 5.7) for players in the $_i$ part of the possible full-path[9], $Prob_j$ is the probability of players following the possible full-path in the split $_j$, n is the total number of parts, and m the total number of splits from the $_i$ part until the end of the possible full-path. When there is no split (i.e., the last part), a probability of 1 is assumed.

$$\text{AvailableContentSplitted} = \sum_{i=1}^{n} AvailableContentPossibleFullPath_i \tag{5.10}$$

Where n is the total number of possible full-paths and $AvailableContentPossibleFullPath_i$ is the available content presented in the possible full-path $_i$ (Equation 5.9 result).

$$\text{MeanAvailableContentSplitted} = (\frac{AvailableContentSplitted}{\sum_{i=1}^{n} PlayersStep_i}) * 100 \tag{5.11}$$

[8] Note that the old Equations 5.7 and 5.8 still valid to compute the available content of not splitted Game Paths.

[9] It is essential to highlight that for each AC computation, only the players on the current part $_i$ are considered, even though the length assumed regards the full possible path and there are players on the other parts. This consideration is needed because the probability must be only applied over the available content of the players that must make a choice, leaving the ones that already did it to be computed further without the influence of this previous probability.

Where $AvailableContentSplitted$ is the result from Equation 5.10, n is the total number of steps, and $PlayersStep_i$ is the number of players in the step $_i$.

Table 48 – Example of the $AvailableContentSplitted$ computation regarding Figure 36

Possible Paths	**Origin**	**Destination**	**Number of Steps**	**Number of Splits**	$\prod_{j=1}^{m} Prob_j$	AC_i	AC_i* ($\prod_{j=1}^{m} Prob_j$)
1	1	1-1-1	7	2	0.91*0.85 (0.78)	2.57	2.02
	1-1	1-1-1	7	1	0.85	0.28	0.23
	1-1-1	-	7	0	1	1.14	1.14
2	1	1-1-2	7	2	0.91*0.14 (0.13)	2.57	0.33
	1-1	1-1-2	7	1	0.14	0.28	0.03
	1-1-2	-	7	0	1	0.14	0.14
3	1	1-2	4	1	0.04	1.5	0.06
	1-2	-	4	0	1	0	0
4	1	1-3	5	1	0.04	2	0.08
	1-3	-	5	0	1	0	0
						Result $\sum$	**4.03**

It is essential to highlight that the proposed metrics are applied from the root segment considering all its variants and not from individual variants, as discarding the original segments' information impairs the notion of "available content". A general metric can be computed summing all the available contents from all common Game Paths (regardless if they are splitted or not), as depicted in Equation 5.12, where $AvailableContentGamePath_i$ regards the available content of the common Game Path $_i$ and n the total number of common Game Paths.

$$\text{GeneralAvailableContent} = \sum_{i=1}^{n} AvailableContentGamePath_i \tag{5.12}$$

When a player follows one variant, he/she leaves the previous segment (i.e., the original segment); thus nothing prevents original segments from becoming empty. Moreover, if a player's first action is not represented by the first step of any common Game Path, a new common Game Path will be built with this action as its first step. This fact highlights an important aspect of the Game Path, the order. A Game Path is similar to another if, and only if, they have the same actions in the same order[10], moreover, the first action of a player will always be the first step of a Game Path, the second action will be the second step, and so on. To better clarify this concept, let us assume the following example presented in Table 49, where four different players (P1, P2, P3, and P4) played a given game with only three available actions (A1, A2, and A3). For each time-span, each player can do actions that change the Game Paths' arrangement. To represent a Game Path

[10] Note that in the proposed method, it is not possible to have two identical Game Paths, instead of it, there will be only one Game Path with two players in the head's position.

textually the following notation is adopted, where each action is linked by an arrow (→), being the action in the left of the arrow the predecessor of the action in the right of it. Assuming a Game Path where the first action is A3 and the second is A1, the corresponding notation will be A3→A1.

Table 49 – Example of Game Paths' constructions according to the influence of the actions' order

	Times-span					
	1st		2nd		3rd	
Players	Action	Resultant Game Path	Action	Resultant Game Path	Action	Resultant Game Path
P1	A1	A1	A2	A1→A2	A1	A1→A2→A1
P2	A2	A2	A2	A2→A2	A3	A2→A2→A3
P3	A1	A1	A3	A1→A3	A2 and A1	A1→A3→A2→A1
P4	A3	A3	-	A3	-	A3

It is possible to see five different aspects of Game Paths in Table 49, where: (1) even though players can do the same actions, the order aspect differs the way that one player plays from the others, entailing in different Game Paths; (2) nothing prevents a player from doing more than one action in a given time-span (like P3 in the 3rd time-span); (3) a player can stop playing (like P4 who played only in the first time-span); (4) if two or more players act identically there will be only one Game Path for all (like P1 and P3 in the first time-span); and (5) the "frequency" idea (like P2 in the first and second time-spans, when he/she did the same action twice consecutively). In the proposed method, it is assumed that the abstract definition of "the sequence of a player's choices" must consider the cases where the same action is performed many times in the same series because it foments this player's behavior, even though it adds complexity to the Game Paths generation process.

During the method implementation in a candidate game, this game could present a high number of complex Game Paths that impairs a desirable monitoring frequency due to one's limited data processing power; therefore, the proposed method must be able to allow some adjustments over the Game Paths complexity to minimize this problem. Although the simplification of Game Paths can be useful, it entails a loss of information. Next, three ways to simplify Game Paths are presented. Note that they can be applied together.

1. To disregard the frequency aspect, considering a series of the same action as only one step of the Game Paths. It reduces the complexity of generating Game Paths because fewer steps would exist; however, the players' behaviors are less detailed.

2. To consider the attained human needs instead of actions. The disadvantage of this approach is that players with different actions regarding the same need will be considered as having the same behavior in the human need perspective, although they behave differently from the action perspective.

3. To establish a desirable size limit of n steps to all Game Paths, keeping only the newest steps and removing the oldest ones that do not have players on them. The disadvantages of this approach are the loss of the players' historical behavior, which impairs the analysis over the patterns of game content consumption, and the possible creation of "new" Game Paths considering the steps removed previously.

Game Paths are hierarchical structures that can increase without a limit; thus, a procedure to segregate analysis according to a desired perspective is needed. Given it, we propose to divide Game Paths by the called "Positions". A Position regards the depth where segments are placed. For example, assuming the Game Path Segment 1-1-4, it is in Position 3, given that it is the third segment from the root segment. Note that the number of Positions in a Game Path structure ranges from one (the root) until positive infinite. Also, by choosing a given Position, a set of Game Path Segments can be referenced.

The specificity degree of a Game Path can be understood in two ways: (1) considering the number of steps of it, and (2) considering the number of splits. Thus, the more steps or splits a Game Path has, the more specific it is, and consequently, the more a player "walks" into this path, the more specific his/her behavior is. Moreover, when a game has Game Paths with few splits, it means that this game is more linear, while when a game has Game Paths with many splits, it is more non-linear. Note that this concept of linear or non-linear is attached to the game content's capacity to attract the players' attention to do different activities. For example, a game designed as non-linear can be seen as linear, in the Game Path perspective, if its contents were not able to please the players enough to encourage them to explore and try new choices.

In conclusion, regardless of the Game Paths complexity configurations, each Game Path represents ways of how players play. In addition to it, each way can be understood as a historical sequence of choices that can be shared between different players, highlighting the idea of chase patterns of human needs. Therefore, each Game Path Segment is assumed as a personality trait, being the number of traits of a game variable according to its players' behaviors. Note that the number of traits linked to a player varies according to its Game Path Segment Position, where the deeper, the more traits are linked to the player. An exciting aspect of Game Paths is that it is possible to identify abandonment reasons of players and associate it to players on the same path because, in the same way that they chase the same human needs or learn similarly, they may abandon the game due to the same reasons[11].

As a final remark, to mathematically formalize the fact that the generated Game Paths regard players' choices (i.e., their personalities) and not a random or restricted

[11] Note that this consideration regards the gameplay aspect, and not external game factors, like the end of a vacation. Thus, this abandonment association approach presents a possible bias related to external game factors.

behavior, the $GeneralPersInfluence$ metric is proposed. This metric encompasses the idea of, given a set of available actions (z) to be performed by any player, what is the probability of a player not performing a sequence of actions based on its own will. As a first step, let us assume the probability of a player performing an action $_i$ randomly, as depicted in Equation 5.13.

$$\mathrm{P}_{\mathrm{random\_i}} = \frac{1}{z} \tag{5.13}$$

However, it is expected that a person's personality reduces the set of possible actions of a game (z) to a value of desirable actions; hence, Equation 5.14 is proposed contemplating a reduction factor, named $pers_{prunning}$.

$$\mathrm{P}_{\mathrm{random\_pruning\_i}} = \frac{1}{z - pers_{pruning}} \tag{5.14}$$

The exact value of $pers_{pruning}$ is unknown, but its range is assumed from 0 until (z -1). A $pers_{pruning}$ value of 0 means that all possible actions are pleasurable to the player, whereas a z - 1 value means that only one action is. As it is assumed that a player plays a game with the hope of finding something desirable or already found it, the resultant of $(z - pers_{pruning})$ will always have a minimum value of 1 (considering that the player wishes to continue playing, even if the only available action is not desirable).

Note that, even though this probability of choosing one action can be minimal (assuming a z of 50 and a $pers_{pruning}$ of 10), it does not depict the fact of players being forced to choose or not one specific action due to the currently available content. To identify such an influence, multiples players' behavior must be observed, highlighting the probability of these players performing the same sequence of actions (i.e., being in the same Game Path Segment). Such as depicted by Equation 5.15, where x is the number of players, and y the number of considered steps.

$$\mathrm{P}_{\mathrm{GamePathSegment\_random\_i}} = \left(\prod_{i=1}^{x} P_{random\_pruning\_i}\right)^{y} \tag{5.15}$$

It is possible to notice that the greater the number of players or the considered steps, the lower is the probability of players doing the same sequence of actions randomly (i.e., being in the same Game Path Segment). Thus, to identify how the players' personalities influence their placement in the same segment, Equation 5.16 is proposed, being it the difference between the random selection of actions based on personality or not.

$$\mathrm{Pers}_{\mathrm{influence\_i}} = \left(\prod_{i=1}^{x} P_{random\_pruning\_i}\right)^{y} - \left(\prod_{i=1}^{x} P_{random\_i}\right)^{y} \tag{5.16}$$

However, as the value of $P_{random\_pruning\_i}$ cannot be identified due to the unreachable $pers_{prunning}$ value (which is specific to each player), the personality influence of players in placing them in the same Game Path Segment is assumed as depicted by Equation 5.17.

$$\text{Pers}_{\text{influence\_i}} = 1 - (\prod_{i=1}^{x} P_{random\_i})^{y} \tag{5.17}$$

A general view of a game considering its players' personalities influence can be obtained by computing the mean $pers_{influence}$ of all Game Paths Segments that contain players, such as depicted in the following Equation 5.18, where n regards the number of Game Path Segments that contain players.

$$\text{GeneralPersInfluence} = (\frac{\sum_{i=1}^{n} Pers_{influence\_i}}{n}) * 100 \tag{5.18}$$

In this case, the nearer 100, the greater is the players' personalities influence in pruning the set of possible actions, whereas, the nearer 0, the lesser is such influence. By comparing an obtained value near 100 with an n value greater than one, it is possible to conclude that players are playing based on what they want to do, as players are following different ways, not being forced to follow a specific sequence of actions.

5.4.2.2 Micro Spectrum Similarity

In the Micro Spectrum, the similarity between players regards the percentage of their sequence of actions that are identical (e.g., the set of previous segments shared by different players that can be placed in different segments). Players in the same Game Path Segment are assumed to have 100% similarity, even though one player performed more actions than others, being placed in a different step of the same segment. This assumption is based on the fact that the opportunity to play for more or less time is not a personality influence.

A game has 100% similarity, in the Micro Spectrum, when all of its players are identified as having the same personality, being all of them placed in the same Game Path Segment. The similarity of a game decreases for every new Game Path created and for every Game Path Segment that is splitted. When a new Game Path is created, the players placed on it do not have anything in common with the players placed on other Game Paths. However, when a Game Path Segment is splitted, the players on the new segments have in common their previous segments. Another remark is that the more players are placed in the same Game Path Segment, the more similar their personalities are. It means that the number of players on each Game Path Segment can be considered as a similarity relevance (i.e., a weight) of that segment compared to the others.

Given all these points, the following Equation 5.19 is proposed to compute the global similarity of each Game Path Segment i:

$$\text{GamePathSegment\_i}_{\text{global\_similarity}} = \frac{\#GamePathSegmentPlayers_i}{\#GamePlayers} \tag{5.19}$$

Where $\#GamePathSegmentPlayers\_i$ regards the number of players that performed the segment i sequence of actions, regardless of these players being placed or not in this segment, and $\#GamePlayers$ is the total number of players of the considered game. Note that the fewer players are counted in a given segment, the less similarity this segment has in the global (game) perspective. It means that for every splitting case, the new splitted segments will have less similarity compared to their previous one.

To compute a game similarity in the Micro Spectrum, Equation 5.20 is proposed based on the idea of a weighted average computation, where, for each Game Path Segment, its value is the result of the Equation 5.19 weighted by the number of players on this segment:

$$\text{MicroSimilarity} = \frac{\sum_{i=1}^{n}(Segment\_i_{global\_similarity} * \#SegmentPlayers_i)}{\sum_{i=1}^{n}\#SegmentPlayers_i} * 100 \tag{5.20}$$

Where n is the number of segments of the considered game, $Segment\_i_{global\_similarity}$ the Equation 5.19 result for the $_i$ segment, and $\#SegmentPlayers_i$ the number of players that was counted in the segment $_i$.

The rationale associated with this computation is that to obtain a similarity degree of a game, each Game Path Segment must have its general behavior percentage weighted by the number of players on that segment. Therefore, the Micro similarity degree of a game regards how familiar or different the players' personalities are regarding their sequence of actions. This metric range is from 0% until 100%, where the nearer 100%, the more similar, whereas the nearer 0%, the more different the players' personalities are. Note that a simple average computation of the $Segment\_i_{global\_similarity}$ of all segments is not applicable as it disregards the weight of each Game Path Segment.

The more players a segment has compared to the others, the more relevant this segment is. When a segment becomes more relevant, it reduces the others' relevance, and consequently, increases or decreases the $MicroSimilarity$ result. Such variability in the $MicroSimilarity$ result depends on whether the Game Paths Segments' values and weights accentuated or attenuated a more common behavior. For example, assuming two Game Paths A and B, where A has 80 players (a more common behavior) and B 20 players (a less common behavior), resulting in a $MicroSimilarity$ of 68%. If A received more ten players (totalizing 90), it would increase the $MicroSimilarity$ to 70.24%; however, if B received more ten players instead of A, it would decrease the *MicroSimilarity* to 60.33%. It means that the more the common behavior is accentuated (i.e., received more players than the other behaviors), the more similar the players' personalities are. In contrast, the more attenuated the common behavior is (i.e., received fewer players than the other behaviors), the more different the players' personalities are. In sum, the *MicroSimilarity* result considers the arrangements of the Game Paths, splitted segments, and the relevance of each segment to provide a similarity measure.

5.4.2.3 Macro Spectrum - The Human Needs Chase Priorities

The Macro Spectrum models the players' personalities based on their chase patterns regarding human needs, differently from the Micro Spectrum that considers the sequence of players' actions. Such chase patterns are observed by the following two perspectives:

1. If a given need is chased or not.
2. What is the priority between the different kinds of needs (i.e., the ranking of human needs more chased).

For example, a player can be unsocial (not chasing the Affiliation need) and prefer to chase the Power need (first choice) instead of the Information one (second choice). Bearing in mind the six types of human needs adopted in this thesis and the possibility of each one being the first, second, third, fourth, fifth, or the sixth most chased need, or not being chased at all, the Macro Spectrum allows the identification of 732 distinct personality traits, where 720 traits regard a different composition of human needs priority chase, and 12 the chase or not of each human need. Tables 50 and 51 present examples of players' Macro personalities. To make these illustrations clearer, Materialism, Power, Affiliation, Achievement, Information, and Sensual are represented by the following respective characters: Mat, Pwr, Aff, Ach, Inf, and Sen.

Table 50 – Example of human needs chase priorities (ranking)

Player ID	**Time-span**	**First need**	**Second need**	**Third need**	**Fourth need**	**Fifth need**	**Sixth need**
1	1st time-span	Aff	Pwr	Mat	Sen	Inf	Ach
1	2nd time-span	Pwr	Aff	Mat	Inf	Sen	Ach
2	1st time-span	Mat	Inf	Ach	Aff	Pwr	Sen
2	2nd time-span	Mat	Inf	Aff	Ach	Pwr	Sen

Table 51 – Example of human needs chase or not

Player ID	Time-span	Mat	Pwr	Aff	Ach	Inf	Sen
1	1st time-span	Yes	Yes	Yes	Yes	Yes	No
1	2nd time-span	Yes	Yes	Yes	Yes	Yes	No
2	1st time-span	Yes	Yes	Yes	Yes	Yes	No
2	2nd time-span	Yes	Yes	Yes	Yes	Yes	Yes

As we can see, Table 50 shows an example of how players can change or not their needs priorities over time, which is the result of two connected things, the players' personality in prioritizing one or another need and the available game content. Complementary, Table 51 depicts, considering the chase priorities of Table 50, the chase or not of human needs (a boolean perspective). Note that a need can be placed in the sixth position being chased, players can have more than one need not chased, and a not chased need can be chased at any moment. It is essential to highlight that these arrangements are linked to the needs offered by a game content. If a need is not present in the content, it cannot be chased.

The success or failure in attaining a human need already portrays a players' desires. Thus, to compute the ranking of human needs, each need attainment attempt is counted, meaning that its numbers of attainments and impairments are summed up. By considering for each time-span the historical sum of each human need attempt, it is possible to identify the order of chase for each need, where the higher the sum, the more prioritize a need is, resulting in a ranking with six positions, such as depicted in Table 50.

An interesting quality of the Macro Spectrum is that it is not susceptible to the frequency aspect as the Micro Spectrum is, as the Macro Spectrum disregard any order aspect, considering only the groups of needs chased or not. Note that few repetitions of the same need attempt would not have enough impact to change the identified traits in the Macro Spectrum perspective.

5.4.2.4 Macro Spectrum Similarity

It is assumed that the similarity between players' personalities in the Macro Spectrum regards how similar the players prioritize their human needs over time. Given this, Equation 5.21 provides a means to convert each player's original Macro Spectrum ranking to a time-based perspective of how much time each need is placed at each position (i.e., first, second, third, fourth, fifth, and sixth) (the called Macro distribution), where i regards a given player, h a given human need, p a possible position, $\#DaysInPosition$

the number of days that the human need h of player i was placed on the position p, and $\#TotalOfDays$ the total number of days that player i played. Tables 52, 53, and 54 present this Equation computation to the hypothetical players A, B, and C.

$$\text{HumanNeedAtPos}_{\text{i\_h\_p}} = \frac{\#DaysInPosition}{\#TotalOfDays} \tag{5.21}$$

Table 52 – Player A Macro distribution

	Positions					
Human Needs	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Materialism	100	0	0	0	0	0
Power	0	100	0	0	0	0
Affiliation	0	0	0	43	57	0
Achievement	0	0	100	0	0	0
Information	0	0	0	57	43	0
Sensual	0	0	0	0	0	100

Table 53 – Player B Macro distribution

	Positions					
Human Needs	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Materialism	100	0	0	0	0	0
Power	0	100	0	0	0	0
Affiliation	0	0	0	0	100	0
Achievement	0	0	50	50	0	0
Information	0	0	50	50	0	0
Sensual	0	0	0	0	0	100

Table 54 – Player C Macro distribution

	Positions					
Human Needs	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Materialism	100	0	0	0	0	0
Power	0	100	0	0	0	0
Affiliation	0	0	40	0	60	0
Achievement	0	0	50	50	0	0
Information	0	0	50	50	0	0
Sensual	0	0	0	0	0	100

Next, the game mean Macro distribution is identified. It regards the computation of the mean value of each human need at each position, such as depicted in Equation 5.22, where i, h, and p have the same meanings of the previous Equation, and n regards the number of players. Table 55 presents the game Macro distribution regarding the players A, B, and C.

$$\mathrm{MeanHumanNeedAtPos_{h\_p}} = \frac{\sum_{i=1}^{n} \mathrm{HumanNeedAtPos_{i\_h\_p}}}{\mathrm{n}} \tag{5.22}$$

Table 55 – Game mean Macro distribution regarding the players A, B, and C

	Positions					
Human Needs	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Materialism	100	0	0	0	0	0
Power	0	100	0	0	0	0
Affiliation	0	0	13.33	14.33	72.33	0
Achievement	0	0	66.66	33.33	0	0
Information	0	0	33.33	52.33	14.33	0
Sensual	0	0	0	0	0	100

Given the mean Macro distribution, it is possible to verify for each player how similar his/her behavior is compared to the game's most common behavior. It can be performed by verifying the absolute difference between a player's human need position value and the equivalent game mean value, such as proposed by Equation 5.23, where $MeanHumanNeedAtPos_{h\_p}$ regards the game mean Macro distribution of human need h at position p, and $HumanNeedAtPos_{i\_h\_p}$ the player i value regarding the same human need h and position p. Tables 56, 57, and 58 present the Macro similarities of players A, B, and C, respectively. Note that a value of 100 is present for all the cases where the players' Macro behavior is identical to the game's most common behavior (i.e., 100% of similarity), being it the placement or the no placement of a need in a given position (the idea of doing or not something with a particular occurrence).

$$\mathrm{Similarity_{i\_h\_p}} = \text{100-ABS}(\mathrm{MeanHumanNeedAtPos_{h\_p}}\text{-}\mathrm{HumanNeedAtPos_{i\_h\_p}}) \tag{5.23}$$

Table 56 – Player A Macro similarity

	Positions					
Human Needs	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Materialism	100	100	100	100	100	100
Power	100	100	100	100	100	100
Affiliation	100	100	86.66	71.33	84.66	100
Achievement	100	100	66.66	66.66	100	100
Information	100	100	66.66	95.33	71.33	100
Sensual	100	100	100	100	100	100

To generate a single similarity value for each player, a mean value of all its similarities values can be computed, such as proposed by Equation 5.24, where its application to

Table 57 – Player B Macro similarity

	Positions					
Human Needs	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Materialism	100	100	100	100	100	100
Power	100	100	100	100	100	100
Affiliation	100	100	86.66	85.66	72.33	100
Achievement	100	100	83.33	83.33	100	100
Information	100	100	83.33	97.66	85.66	100
Sensual	100	100	100	100	100	100

Table 58 – Player C Macro similarity

	Positions					
Human Needs	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Materialism	100	100	100	100	100	100
Power	100	100	100	100	100	100
Affiliation	100	100	73.33	85.66	87.66	100
Achievement	100	100	83.33	83.33	100	100
Information	100	100	83.33	97.66	85.66	100
Sensual	100	100	100	100	100	100

players A, B, and C values generates the following similarity values, respectively: 94.7, 96.61, and 96.66.

$$\mathrm{PlayerMacroSimilarity_i} = \sum_{\mathrm{h=1}}^{6} \sum_{\mathrm{p=1}}^{6} \mathrm{Similarity_{i\_h\_p}} \tag{5.24}$$

As a final procedure to identify a game's Macro similarity, a mean value over all its players' similarities can be computed, such as suggested by Equation 5.25 (where n is the number of players), entailing in the value 95.99 regarding players A, B, and C. It means that, in the hypothetical game, its players prioritize the same human needs with a similarity of 95.99%.

$$\mathrm{MacroSimilarity} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{PlayerMacroSimilarity_i}}{\mathrm{n}} \tag{5.25}$$

In sum, the Macro similarity of a game can be computed based on the following four steps:

1. To compute to each player his/her Macro distribution.
2. To compute the game mean Macro distribution based on all players' Macro distributions.
3. To compute each player's similarity between his/her Macro distribution and the game mean Macro distribution.

4. To compute the mean of all players' similarities as the game Macro Similarity value.

As a final remark, note that this computation should be performed to each time-span to highlight changes in players' Macro similarity over time.

5.4.3 Emotion Identification Process

The emotion identification process has two approaches: one for the internal emotions (Section 5.4.3.1) and another for the social ones (Section 5.4.3.2). However, regardless of the emotion type, all emotions are identified based on the attainment or impairment of human needs.

5.4.3.1 Internal Emotions

The internal emotion approach demands three kinds of information: (1) the player's goal, (2) its likelihood, and (3) the final result. However, there are some cases where the final result is not needed. Next, in Table 59, all internal emotions considered in this work are presented together with their requirements and the respective polarity.

Table 59 – Internal emotions' requirements

Emotion	**Goal**	**Likelihood**	**Final Result**	**Polarity**
Hope	X	X		Positive
Fear	X	X		Negative
Joy	X	X	X	Positive
Distress	X	X	X	Negative
Satisfaction	X	X	X	Positive
Fears-confirmed	X	X	X	Negative
Disappointment	X	X	X	Negative
Relief	X	X	X	Positive

The Hope and Fear emotions are related to the situation where one does not know about the final result. Joy and Distress happen in situations where one knows the final result of something, even though it has not happened yet. Satisfaction, Fears-confirmed, Disappointment, and Relief occur according to a previous emotion (Hope or Fear). For more details about the occurrence situations of each internal emotion please see the Sections 2.6.3 and 2.7.

As emotions are considered a short-term psychological aspect (shorter than sentiment and personality) (CARVER; SCHEIER, 2012), their identification is made based on the shortest time unit adopted in the proposed method, the time-span. For each time-span, a summary of human needs attained, impaired, and not chased is considered. Let us assume the representation depicted in Figure 37, where below each human need group, the "+" sign represents an attainment, and the "-" sign an impairment ("None" is used when the human need was not chased). Moreover, when a human need has a positive value (i.e., the number of attainments is greater than the number of impairments), it receives the green color, when it has a negative value a red color, and when the number of attainments and impairments are equal or when the need was not chased, the gray color is adopted.

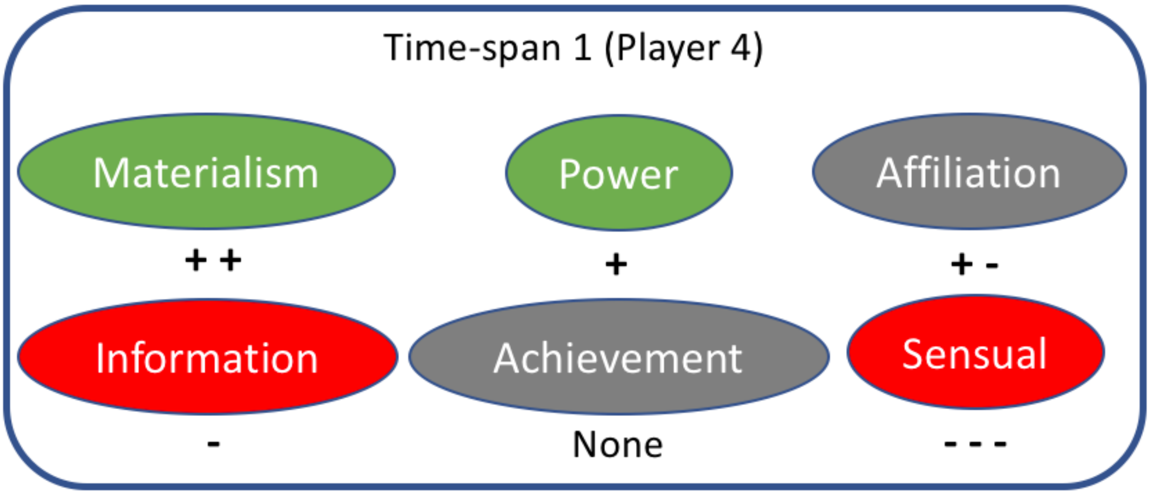


Figure 37 – Example of attainment, impairment, and not chase of human needs

It is assumed that an impairment or attainment of a human need represents the chase of a goal, regardless of achieving it (attaining) or not (impairing). Therefore, the identification of the players' goals is made individually and a posteriori of each time-span.

The likelihood and the final result identifications are made comparing two time-spans, the last one (n) and its previous ($n-1$). When a human need has a green color in the previous time-span, it means a probability of more than 50% of attaining this need in the next time-span (generating a Hope emotion), when it has a red color, it means less than 50% of chance to attain such need (generating a Fear emotion), and when the color is gray, it means exactly 50% of chance (note that this value is assumed to situations where the need has never been chased; the Hope emotion is also assumed in this case). In addition to it, the final result is considered as being each attainment and impairment presented in the time-span n. Table 60 summarizes the identification rules for the internal emotions according to the time-spans configurations.

Special consideration is given to the identification of Hope and Fear. These emotions mean, respectively, a positive and negative prospect about attaining needs in the next time-span. However, their identification cannot consider their previous occurrences. It is justified because they are based on the emotions entailed by final results and not on their expectancy. To clarify, let us assume the following example where a given player has Hope and Disappointment in a given time-span. By considering the balance of 50% between positives and negatives emotions, the Hope emotion would be generated for the next time-span; however, the correct emotion should be Fear because the only attempt was

Table 60 – Summary of internal emotions identification rules in the proposed method

Emotion	Time-span $n-1$ Situation	Time-span n Situation
Hope	A green color.	-
Fear	A red color.	-
Joy	A green color (with 100% of attainment).	A green color (with 100% of attainment).
Distress	A red color (with 100% of impairment).	A red color (with 100% of impairment).
Satisfaction	A green color.	A "+" sign.
Fears-confirmed	A red color.	A "-" sign.
Disappointment	A green color.	A "-" sign.
Relief	A red color.	A "+" sign.

not successful. Given this, Hope and Fear's computations only consider the occurrences of Joy, Distress, Satisfaction, Disappointment, Fears-confirmed, and Relief[12].

For each attainment or impairment of the time-span n, an emotion will be identified according to the rules presented in Table 60. To better clarify this identification process, let us assume the example depicted in Figure 38, where only the Materialism need is considered through eight consecutive time-spans. Moreover, Table 61 points the emotions identified to each consecutive pair of time-spans of Figure 38. The signs "+*" and "-*" are adopted to reference the Hope and Fear occurrences, respectively.

Table 61 – Internal emotions identified in Figure 38

Time-span Number $(n-1)$	Time-span Number (n)	Emotions Identified in Time-span n
1	2	Hope, Satisfaction, and Joy
2	3	Hope, and Disappointment
3	4	Fear, and Fears-confirmed
4	5	Fear, Fears-confirmed, and Distress
5	6	Fear, and Relief
6	7	Hope, and Disappointment
7	8	Fear, and Relief

Note that no emotion can be identified in the first time-span, as there is no information about the likelihood.

5.4.3.2 Social Emotions

The social emotion approach requires four kinds of information: (1) the causer, (2) the affected, (3) if one likes or not the causer or affected, and (4) if the event was

[12] Note that the number of attainments and impairments of a given time-span can also be used to identify the Hope and Fear emotions.

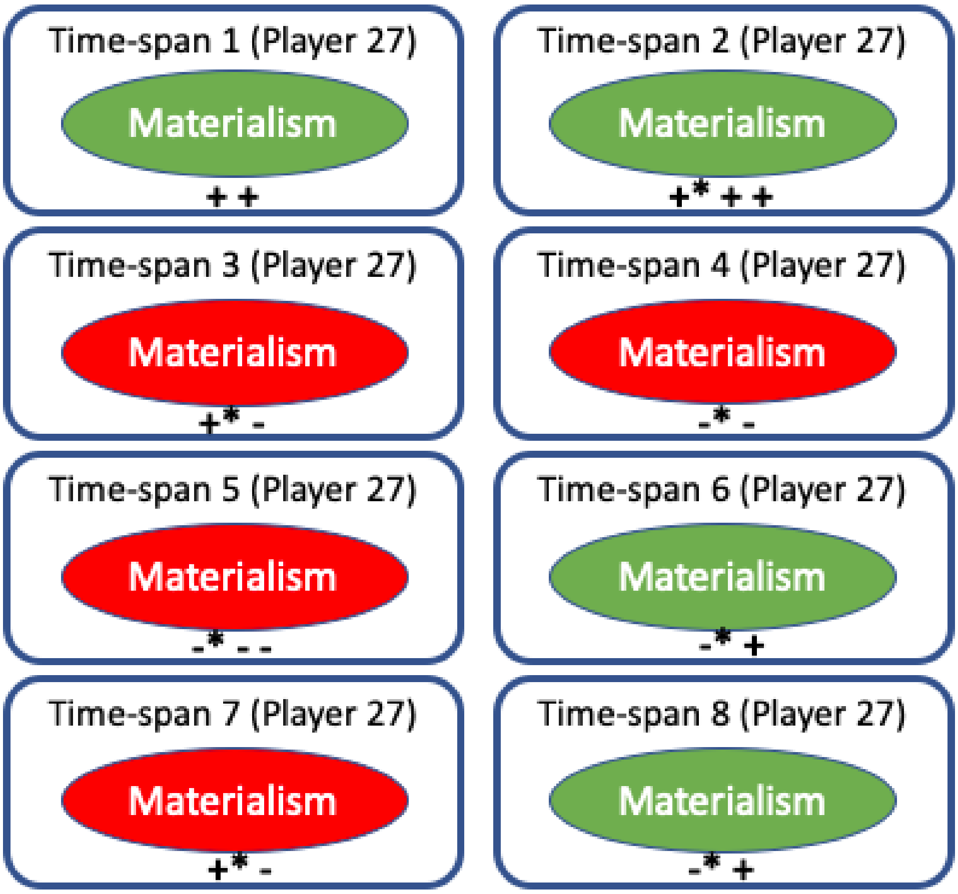


Figure 38 – Sequence of attainment and impairment of the Materialism need

desirable or not. However, the "like" aspect is not needed in some cases. Table 62 depicts the conditions to each social emotion happen together with its respective polarity.

Table 62 – Social emotions' requirements

Emotion	**Liked NPC or Player**	**Causer**	**Affected**	**Desirable Event Self**	**Desirable Event Other**	**Polarity**
Anger	-	Other	Self	No	-	Negative
Shame	-	Self	Other	No	No	Negative
Gratitude	-	Other	Self	Yes	Yes	Positive
Gratification	-	Self	Other	Yes	Yes	Positive
Happy-for	Yes	Other	Other	Yes	Yes	Positive
Pity	Yes	Other	Other	No	No	Negative
Gloating	No	Other	Other	Yes	No	Positive
Resentment	No	Other	Other	No	Yes	Negative

It is interesting to highlight that a social emotion's polarity can change to the same circumstance depending on whether the affected person is someone that one likes or not (e.g., Pity and Gloating). Moreover, even though there are roles regarding causer and affected, the final polarity considers only the self, regardless if he/she is the causer or the affected. As an additional remark, the social emotions do not consider the expectancy that something good or not happen to others (such as the Hope emotion); it only considers the event's final result.

From the self point of view (i.e., a given player), the definition of desirable and undesirable events can be represented in two ways, one regarding the attainments or the impairments of a self's human needs, and the other considering the events that happened toward others. Even though some events do not affect the self directly, his/her Affiliation need may be attained or impaired due to the nurturance aspect of it, which regards the wish that good things happen to others (see Section 2.7 for more details)[13], therefore, it is also possible to attain or impair the Affiliation need even when events affected others (whom the self likes or hates). It is essential to highlight that for social emotions where the self is affected, any human need group can be influenced. In conclusion, when an event happens to a player (self) or another player or an NPC that the player likes or dislikes, a social emotion can be identified since the causer, and the affected are known. Note that, except for the Anger, Shame, Gratitude, and Gratification emotions, no social emotions are identified when the self does not know (i.e., like or hate) the other. Moreover, a special consideration regards the acknowledgment of the self about the occurrence of an event, as even though the self likes or dislikes someone and an event happened to him/her, no social emotions can be identified if there is no awareness of it by the self. Therefore, the usage data must specify when one knows about the occurrence of an event or not.

Although the identification of social emotions differs from the one for internal emotions, all social emotions are linked to the attainments or not of self's needs, and hence, they are linked to internal emotions. For example, when a player obtains an item from another player, he/she attains a Materialism need, having Satisfaction or Relief of obtaining it and Gratitude toward the other; otherwise, if the item was obtained only by the self's effort, there is no Gratitude. Therefore, it is assumed that the occurrence of a social emotion entails the occurrence of an internal one, also influencing its likelihood. Table 63 points some examples of social emotions identified together with internal ones in a given time-span (when the internal emotion is attached to a social one, the "*" sign is used). Note that the Hope and Fear emotions are affected indirectly by the presence of social emotions, as they are defined by the number of positive and negative emotions (internal and social) in the previous time-span, while the other internal emotions can be directly attached to social ones in a relation of one to one. Moreover, it is considered that a player likes or hates another one or an NPC depending on whether the number of positive social emotions are greater than the negative ones (i.e., like) or not (i.e., hate) regarding this NPC or player.

It is essential to highlight that the attached social and internal emotions always

[13] Although the nurturance aspect does not regard the wish that bad things happen to others, we opted to link this situation to the Affiliation need because it represents a possible social interaction.

Table 63 – Example of internal and social emotions occurrences

	Time-span n	
	Internal Emotion	**Social Emotion**
Materialism	Hope, Satisfaction*, and Satisfaction	Gratitude
Power	Hope and Disappointment*	Anger
Affiliation	Fear, Fears-confirmed*, Fears-confirmed*, Relief*, and Relief	Pity, Resentment, and Happy-for
Achievement	None	None
Information	None	None
Sensual	Fear, and Relief	None

have the same polarity. Thus, when a human need is attained due to a social aspect, it weights double compared to an attainment without it in the Sentiments computations (Equations 5.1 and 5.2), as there are two emotions instead of only one. We opted to apply this "double weight" because social interaction is considered an important motivational factor to play games, according to the conclusions of Section 2.6.3.

5.5 Method Assessment

The proposed method was conjectured based on the "Aggregated psychological essence" property of the unified human-being model, presented in Section 5.1, being it posteriorly formalized in Figure 26 (the method fundamental structure). Even though the overall method is incipient, each proposed link regarding its psychological aspects has references that foment them, as summarized in Table 64.

Besides the academic references, we enjoyed conceptually assessing the proposed method with two psychologists, a professor, and a researcher. This assessment process consisted of the author of this thesis presenting to them his interpretations about the psychological aspects and the proposed connections between them. The experts checked each interpretation by pointing the flawed understandings and providing the correct interpretation and validating the proposed philosophical connections between these aspects. Besides, additional literature was suggested to support a better understanding of the fundamental pillars of psychology. Both stated that the method is coherent and confirmed our answers to RQs 1, 2, 4, 5, 6, and 11 (the ones that regard pure psychological aspects, without the game context; except by RQ 11, which is answered next).

[14] Examples of personality tests and their references can be found at <https://www.sigmaassessmentsystems.com/assessments-category/personality-tests/>.

Table 64 – References to each link of the proposed method

Link	References
Action to Human Need	(MURRAY, 1938), (SWEETSER et al., 2003), (BOSTAN, 2009),(WEILLER, 2015), (MASLOW, 1968), (CZISIKSZENTMIHALYI, 1990), (DECI; RYAN, 1985), (DECI; RYAN, 1995), and (SCHRAMM; LYLE; EDWIN, 1961)
Human Need to Emotion	(SCHERER; SCHORR; JOHNSTONE, 2001), (SCHERER, 1987), (SCHERER, 2001), (BOSTAN, 2009), (CZISIKSZENTMIHALYI, 1990), and (SCHRAMM; LYLE; EDWIN, 1961)
Emotion to Sentiment	(MUNEZERO et al., 2014), (FRENCH, 1947), and (PLUTCHIK, 1980)
Human Need to Personality Trait	All references regarding the questionnaires applied to identify personality traits. The total number of references to all personality models is unknown[14].

5.6 Answers to RQs 10 and 11

The answer to RQ 10 “To what extent characteristics of usage data can be used to identify psychological profiles?” is based on the psychological profile proposition, depicted in Section 5.1. After answering the RQs 1 until 9, it was possible to understand psychological aspects, psychological models, and their linkage, where two unified models were obtained from the proposed SLR through two applications of the Unification Explorer Framework. As a result, the proposed method was conjectured based on the unified human-being model’s simplicity and coverage characteristics, where a set of psychological aspects are identified based on the players’ actions in-game. The proposed psychological profile is then comprehended as a set of metrics that contains psychological descriptions linked to it. Until the present moment, and as far as our knowledge and efforts allowed us, the extent of psychological profile identification in usage data regards emotions, sentiments, human needs, and personality traits, representing only 40.59% of the 101 GCs of the unified human-being model. The identification of the remaining 59.41% of the general human-being model’s GCs and the 100% of the general player model’s GCs are left for future works, where their findings can expand the extent to which characteristics of usage data can be used to identify improved psychological profiles of players.

The answer for RQ 11 “How an identified profile on usage data can be assessed?” is based on the positive and negative aspects of the proposed method, as depicted in Section 5.2. On the one hand, it is impossible to assume the identified psychological aspects as truthful because it is impossible to assess them with an expert considering a massive number of players or verify them with each player; thus, the identified psychological aspects are considered theoretical. However, on the other hand, the identified aspects can be used to improve risk prediction models, explain players' past behavior, and identify possible future trends, highlighting risk situations before they happen. Bearing in mind it, an additional assessment is proposed, which regards the use of the generated psychological profile of players in a churn prediction problem, where if the psychological features present a similar or better performance than the state-of-the-art approach, the profile is assessed as accurate, at least, to the churn prediction problem. This assessment is presented in Chapter 7. Despite assessments complexities, it is possible to take advantage of the generated psychological profile of players to improve the understandings of the players' enjoyment of the game in short, mid, and long-term perspectives; through the analysis of the attained or impaired human needs (short-term), the generated emotions (short-term), the historical changes of sentiments (mid-term), and the identified personality traits (long-term).

6 Method Application, Analysis, Discussions, and Comparison

This Chapter describes the proposed method application regarding its configurations (Section 6.1) and technical aspects (Section 6.2), the experimental results, analysis, and discussions regarding the identified psychological profile of players (Section 6.3), and a comparison between the proposed method with the Game Refinement Theory (Section 6.4), highlighting how the proposed method relates to a parallel approach that has the same aim, a better comprehension of players.

6.1 Method Configuration

The proposed method was implemented into a system named "**P**layer **P**sychological **P**rofile Identification System", or just 3PIS, and applied to the MMORPG Blade&Soul (Figures 39 and 40 show the game theme and gameplay respectively), an entertainment-focused game that contemplates all the method's essential assumptions depicted in Section 5.2. Its usage data was obtained from a Data Mining Competition (LEE et al., 2018) and has a total of 10,000 players[1] divided into three subsets (as some players are present in more than one subset, the number of unique players is 9,647), containing a total of 23 weeks (155 days). Table 65 summarizes some dataset aspects, such as the period of data collection, the number of players, the number of instances (each instance refers to a player's action), and the business model adopted. Note that there are some gaps between the period of each subset; moreover, the business model changed in the third subset[2]. As an additional remark, the subsets' names are "Training", "Test_1", and "Test_2" (these names are linked to the Data Mining Competition, and their meanings are detailed in Chapter 7, where the method is assessed in a churn prediction problem). It is essential to not misunderstand these names with the ones of the Data Mining process.

Before computing the psychological profile for the Blade&Soul, it is needed to analyze its usage data and propose an "action-need map" based on it, as depicted in Section 6.1.1. It is essential to highlight that the action-need map generation is the only manual step of the method.

1 This dataset contains only a sample of the total of active users. Also, all selected players are considered profitable. In the industry colloquial language, these players are known as "whales" (MCALOON, 2018a).

2 The implications of this change are considered in Sections 6.3, and 6.4.1.

Figure 39 – Blade&Soul has an Asian theme, extracted from <https://www.bladeandsoul.com>

Figure 40 – Blade&Soul gameplay, extracted from <http://news.mmosite.com>

6.1.1 Data Preprocessing and Action-need Map Generation

After collecting the data, it is needed to identify and understand the meaning of each possible action in-game and then manually associate it to the attainment or impairment of human needs (i.e., building the action-need map). This process was performed based on the definitions depicted by (BOSTAN, 2009) and summarized in Section 2.7. The dataset presents 82 original actions that attain or impair the needs for Materialism, Power, Affiliation, Achievement, and Information. However, a preprocessing was needed to distinguish when a player wins a PvP battle or not and also to identify the success or failure of an item transformation; thus, the final number of considered actions is 87, due to

Table 65 – Summary of the Blade&Soul dataset

Subset	**Number of Players**	**Number of Instances**	**Business Model**	**Period of Collect**	**Number of Weeks**
Training	4,000	175,139,564	Subscription	2016-03-30 until 2016-05-10	6
Test_1	3,000	197,661,989	Subscription	2016-07-13 until 2016-09-13	9
Test_2	3,000	206,758,995	Free-to-play	2016-12-14 until 2017-02-07	8

the splitting of four PvP actions into winning and losing, and the result of a transformation attempt into success or failure[3]. The action-need map for the Blade&Soul dataset is shown in Table 66, where the information inside the parentheses regards the specific characteristic of the human need that allowed its association, as depicted by (BOSTAN, 2009). Note that when an action is not linked to any human need, the "-" sign is used; moreover, the "*" sign is used when there is a special consideration about a given action (the considerations are described further). As an additional remark, the dataset does not present the information of causer and affected; therefore, it is impossible to compute the social emotions, being only possible to compute the internal ones.

Table 66 – The action-need map for the Blade&Soul dataset

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
EnterWorld	When a player enters the game.	-	-
LeaveWorld	When a player leaves the game.	-	-
EnterZone	When a player enters a zone (a map).	-	-
LeaveZone	When a player leaves a zone (a map).	-	-

To be continued

[3] Note that, sometimes, the actions presented in the usage data can be seen as results of events.

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
Teleport	When a player uses a teleport to change his/her zone.	-	-
DeletePC	When a player deletes his/her avatar.	-	-
PcLevelUp	When a player levels up his/her avatar level.	Achievement (Achievement)	-
GetExperience*	When a player obtains experience points.	-	-
GetMoney	When a player obtains in-game money.	Materialism (Objects Acquisition)	-
SpendMoney	When a player spends in-game money.	-	Materialism (Objects Acquisition)
GetItem	When a player obtains a new item.	Materialism (Objects Acquisition)	-
LoseItem	When a player loses an item.	-	Materialism (Objects Acquisition)
InviteParty	When a player invites another player to join his/her party.	Power (Deference, Dominance), Affiliation (Affiliation, Succorance), Information (Exposition)	-
JoinParty	When a player joins another player's party.	Power (Deference, Dominance), Affiliation (Affiliation, Nurturance), Information (Exposition)	-

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
RefuseParty	When a player refuses to join another player's party.	Affiliation (Rejection)	Power (Deference, Dominance), Affiliation (Affiliation, Nurturance), Information (Exposition)
DismissParty	When a player leaves his/her party.	Affiliation (Rejection)	Power (Deference, Dominance), Affiliation (Affiliation, Succorance, Nurturance), Information (Exposition)
KickParty Member	When a player removes one player of his/her party.	Affiliation (Rejection)	-
Exhaustion	When a player loses his/her HP but still alive to try to save his/her life.	Achievement (Harm Avoidance)	Power (Aggression)
Die	When an exhausted player dies.	-	Achievement (Harm Avoidance), Power (Aggression)
Resurrect	When a player becomes alive again after his/her last death.	-	-
KillNPC	When a player kills an NPC.	Power (Aggression)	-
KillPC	When a player kills another player.	Power (Aggression)	-
DuelEnd (PC) Win*	When a player wins an individual duel against another player.	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition)	-

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
DuelEnd (PC)Lose*	When a player loses an individual duel against another player.	-	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition)
DuelEnd (Team)Win	When the player is the leader of a team and it wins a duel against another team.	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition), Affiliation (Nurturance)	-
DuelEnd (Team)Lose	When the player is the leader of a team and it loses a duel against another team.	-	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition), Affiliation (Nurturance)
MoveToArena	When a player goes to the arena (a PvP place).	-	-
PartyBattleEnd (Team) Win	When the player is the leader of a party and it wins a duel against another party.	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition), Affiliation (Nurturance)	-
PartyBattleEnd (Team) Lose	When the player is the leader of a party and it loses a duel against another party.	-	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition), Affiliation (Nurturance)

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
PartyBattle Result(PC) Win*	When a player is in a party and wins a battle against another player from another party.	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition), Affiliation (Nurturance)	-
PartyBattle Result(PC) Lose*	When a player is in a party and loses a battle against another player from another party.	-	Power (Aggression), Achievement (Harm Avoidance, Recognition, Exhibition), Affiliation (Nurturance)
OccupyBase	When a player occupies an enemy base.	Materialism (Objects Acquisition), Achievement (Recognition, Exhibition)	-
LootItem*	When a player gets an item from a dead NPC.	-	-
UseItem	When a player uses an item.	-	-
DestroyItem*	When a player destroys (i.e., loses) an item.	-	-
GetLootMoney*	When a player obtains in-game money from a dead NPC.	-	-
PartyAuction Start	When a player of a party starts an auction regarding his/her obtained item.	-	-

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
BidParty Auction	When a player bids an auction item.	-	-
PartyAuction Success*	When a player of a party successfully buys an auction item that was obtained by this party.	-	-
Distribute AuctionMoney*	When a player in a party sells an item in an auction and the sale money is distributed towards the party's members.	Affiliation (Nurturance)	-
UnEquipItem	When a player unequip an item.	-	-
SaveEquipInfo	When the information of a player's equipment is saved. Some situations when it happens are during a resurrection or the acquirement of a quest.	-	-

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
RevealItem	When a player identifies an item, discovering its characteristics.	Information (Cognizance)	-
ConsumeGem ByReveal*	When a player uses gems to reveal an item.	-	-
GetItemBy Decomposition*	When a player decomposes an item to obtain its materials (items).	-	-
Expand Warehouse	When a player expands his/her warehouse.	Materialism (Objects Order)	-
RepairItem	When a player repairs an item to avoid losing it.	Materialism (Possessions Retention)	-
GrowUpItem	When a player evolves an item.	Materialism (Construction of Objects)	-
ResultOf Transform Success*	When a player successfully transforms an item.	Materialism (Construction of Objects)	-
ResultOf Transform Failure*	When a player is unsuccessful in transforming an item.	-	-
ExceedItem Limit	When a player reaches the quantity limit of a given item.	-	-
ChangeItem Look	When a player changes the appearance of an item.	Achievement (Autonomy)	-

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
FeedingResult*	When a player feeds his/her pet, obtaining improvement of his/her avatar's attributes for a period of time.	-	-
TradeGiveItem*	When a player gives an item to another player.	Affiliation (Nurturance)	-
TradeGetItem*	When a player receives an item from another player.	Affiliation (Succorance)	-
SellItem*	When a player sells an item to an NPC.	-	-
BuyMyItem*	When a player buys an item from an NPC.	-	-
TradeGive Money*	When a player gives money to another one.	Affiliation (Nurturance)	-
TradeGet Money*	When a player receives money from another player.	Affiliation (Succorance)	-
GetItemFrom NPC*	When a player receives an item from an NPC.	Affiliation (Succorance)	-
DepositItem	When a player puts an item into a warehouse (a safe-deposit).	Materialism (Possessions Retention)	-

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
RetriveItem	When a player retrieves an item from the warehouse.	-	Materialism (Possessions Retention)
PutMain Auction	When a player puts an item for bidding.	-	-
BuyItemNow MainAuction*	When a player buys an auction item.	-	-
UseGathering Item	When a player uses a gathering item.	-	-
GetGathering Item*	When a player gets a gathering item.	-	-
ExpireEvent Item	When an item in a player's inventory expires (i.e., it loses its functionality).	-	-
AcquireSkill*	When a player acquires a new skill.	Achievement (Achievement)	-
SkillLevelUp	When a player's skill levels up.	Achievement (Achievement)	-
LearnTraining	When a player learns a new training that can improve his/her already acquired skills.	Achievement (Achievement)	-
AcquireQuest	When a player acquires a quest.	-	-
CompleteQuest	When a player completes a quest.	Information (Cognizance, Understanding)	-
DisposeQuest	When a player gave up a quest.	-	Information (Cognizance, Understanding)

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
GetQuestItem*	When a player receives an item linked to the quest story.	-	-
GetQuestSkill*	When a player acquires a new skill related to a quest.	Achievement (Achievement)	-
GetChallenge TodayItem*	When a player receives an item related to a daily challenge.	-	-
Complete ChallengeToday	When a player completes a daily challenge.	Achievement (Recognition)	-
GetChallenge WeekItem*	When a player receives an item related to a weekly challenge.	-	-
Complete ChallengeWeek*	When a player completes a weekly challenge.	-	-
CreateGuild*	When a player creates a guild.	-	-
DestoryGuild*	When a player destroys his/her guild.	-	-
GuildLevelUp	When a player's guild levels up.	Achievement (Achievement), Affiliation (Nurturance)	-
InviteGuild	When a player invites another player to join his/her guild.	Power (Deference, Dominance), Affiliation (Affiliation, Succorance), Information (Exposition)	-

To be continued

Action	**Action's Description**	**Human needs Attained**	**Human needs Impaired**
JoinGuild	When a player joins another player's guild.	Power (Deference, Dominance), Affiliation (Affiliation, Nurturance), Information (Exposition)	-
RefuseGuild Invite	When a player refuses to join another player's guild.	Affiliation (Rejection)	Power (Deference, Dominance), Affiliation (Affiliation, Nurturance), Information (Exposition)
DissmissGuild	When a player leaves his/her guild.	Affiliation (Rejection)	Power (Deference, Dominance), Affiliation (Affiliation, Succorance, Nurturance), Information (Exposition)
KickGuild Member	When a player removes one player of his/her guild.	Affiliation (Rejection)	-

In accordance with the data generation process, it was needed to take into account some special considerations over some actions (the ones highlighted by the "*" sign). For example, two actions can be generated in the usage data for the same single event, demanding special attention to identify case-by-case when the doubled actions are accepted or not as two attainments or impairments of human needs.

- The "GetExperience" action happens when a player obtains experience (e.g., killing a monster or completing a quest), being it linked to the attainment of the Achievement need. Moreover, the experience points are used to define when a player levels up (i.e., the "PcLevelUp" action). However, there are situations where the obtained points do not benefit the player, for example, when he/she is already at the maximum level. Therefore, to only consider the attainment of the Achievement need when there is some benefit to the player, only the "PcLevelUp" action is linked to it. As an additional remark, the experience points acquisition can be a consequence of players' actions that do not focus on it, such as when they kill monsters to loot their items, entailing in acquiring experience points by doing so.

- The actions "DuelEnd(PC) Win", "DuelEnd(PC) Lose", "PartyBattleResult(PC) Win", and "PartyBattleResult(PC) Lose" regard the fact of a player winning (a Power attainment) or losing (a Power impairment) a PvP battle against one opponent. In addition to it, there are two victory conditions, (1) when the player kills the other player, and (2) when the duel time is over, and the player did more damage than the other. On the one hand, when a player wins a battle killing the other player, the action "KillPC" is generated by the game, entailing two occurrences of the Power attainment (as "KillPC" also attains this need). On the other hand, when a player is killed by the opponent, the action "Die" is generated, entailing a double impairment of Power. Assuming that players always aim at killing the other, the final result of a PvP battle can have two interpretations, one where a player has a "strong" victory (killing the other) and another where it does not happen, a "simple" victory. Thus, to differentiate the occurrences of strong and simple victories (and also defeats), the links between these actions to the Power need were kept, entailing in double attainment or impairment in case of a strong victory or defeat, and single attainment or impairment in case of a simple victory or defeat.

- The "LootItem" action happens when a player loots a dead NPC looking for items (a Materialism attainment). However, as the "GetItem" action is always generated when an item is looted (also a Materialism attainment), it was opted not to associate any human need to this action, as the event of acquiring an item is already represented by another action.

- The "DestroyItem" action occurs when a player intentionally destroys his/her item. Moreover, the game always attaches to this event the "LoseItem" action; therefore, there is no need to associate this action to any human need because another action already represents this event.

- The "GetLootMoney" action regards a player acquiring money after looting a dead NPC. However, as the game always generates the "GetMoney" action linked to this one, no human needs were associated with this action, as the occurrence of another action already comprehends the event.

- The "PartyAuctionSuccess" action regards a player winning an auction, which entails acquiring an item by this player. Furthermore, as similar to the "LootItem" case, the game always generates a "GetItem" action linked to this action; therefore, it was opted to not assign any human need to it to keep the number of attainments fair to the number of events.

- The "DistributeAuctionMoney" action regards the sharing of a sale between a party's members. However, the game always associates to it the "GetMoney" action; thus, it was opted to associate only the Affiliation need to this action, as the money acquisition (Materialism) is already represented by another action, missing only the social aspect to be represented.

- The “ConsumeGemByReveal” action regards the use of game items (“gems”) to identify an item. However, as this identification is represented by the “RevealItem” action and the gems consumption by the “LoseItem” action, no needs are linked to this action.

- The “GetItemByDecomposition” action happens when a player decomposes an item to obtain its materials (i.e., items). However, as the game attaches to this process the actions “LoseItem” and “GetItem”, there is no need to link this action to any human need because other actions already depict its attainment and impairment.

- The “ResultOfTransform Success” and “ResultOfTransform Failure” actions regard the success or failure of an item transformation. However, regardless of whether it was successful or not, the “LoseItem” action is always attached to this event, whereas the “GetItem” action only occurs when the transformation was successful. Nevertheless, despite the possibility of getting or losing an item, a successful transformation is an attainment by itself; thus, this action (“ResultOfTransform Success”) is linked to the Materialism need, regardless of the attached actions. In contrast, the “ResultOfTransform Failure” action is not linked to any human need, as the “LoseItem” already represents this event.

- The “FeedingResult” action regards the fact of a player feeding his/her pet and then obtaining temporary improvements of his/her avatar’s attributes. This fact can be seen as an Achievement attainment, as the player is improving his/her abilities, but as the game does not give the information of when the improvement expires, the correspondent impairment cannot be identified. Thus, to keep the Achievement need identification balanced, no human need is linked to this action.

- The “TradeGiveItem” action happens when a player gives an item to another one. On the one hand, the player loses the item (which is already represented by the “LoseItem” action; a Materialism impairment), and on the other hand, he/she attains the Affiliation need by doing so. Therefore, as the attached action already comprehends the loss, only the attainment of the Affiliation need is linked to this action.

- Similar to the situation depicted for the “TradeGiveItem” action, the “TradeGetItem” action regards the acquisition of an item, but as the “GetItem” is already attached, only the Affiliation need is considered to this action.

- The “SellItem” action regards the fact of a player selling an item to an NPC, and attached to it, the occurrence of the “GetMoney” and “LoseItem” actions. Thus, as other actions already represent this event, it is not needed to link this action to any human need.

- Like the “SellItem” case, the “BuyMyItem” action comprehends a purchase of an item by a player, being it attached to the occurrences of “SpendMoney” and “GetItem” actions. Therefore, due to the attached actions, there is no need to link this action to any human need.
- The “TradeGiveMoney” action is similar to the “TradeGiveItem” one, being the only difference the gift type. Moreover, as the “SpendMoney” action is always attached to this one, only the link to the Affiliation need is kept.
- Similarly to the “TradeGetItem” action, the “TradeGetMoney” action happens when a player receives money from another player, and as it is always attached to the “GetMoney” action, only the Affiliation aspect is considered in this case.
- The “GetItemFromNPC” action regards the fact of a player receiving an item from an NPC, and as it is always attached to the “GetItem” action, only the Affiliation need is linked to it.
- The “BuyItemNowMainAuction” action occurs when a player buys an item in an auction. However, as it is always attached to this event a “GetItem” action, there is no need to relate it to any human needs.
- The “GetGatheringItem” action happens when a player gets a gathering item, and as always, the “GetItem” action is attached to this one, there is no need to associate it to any human need.
- The “GetQuestItem” action regards the fact of a player receiving an item linked to a quest; however, as the “GetItem” action is always attached to it, there is no need to relate it to any human need.
- The “AcquireSkill” and “GetQuestSkill” actions regard the same situation (a skill acquisition), but both are linked to the Achievement need as they do not happen together.
- The “GetChallengeTodayItem” action happens when a player receives an item associated with a daily challenge, but as the “GetItem” action is always attached to it, there is no need to link it to any human need.
- The “GetChallengeWeekItem” and “CompleteChallengeWeek” actions do not have instances in the usage data; therefore, no needs are linked to them.
- The “CreateGuild” and “DestroyGuild” actions regard the creation and destruction of a guild. In addition to it, when a guild is created, it has only one player (its

owner); moreover, for a guild to be destroyed, only its leader has to be into it. Thus, as these actions do not have any social interaction, no human needs are related to them.

Table 67 summarizes the number of actions that attains and impairs each human need. As we can see, except for the Power need that has the same number of attainments and impairments, the game content provides more opportunities for players to attain their needs than impair them. Moreover, the human need most offered is the Affiliation one, highlighting the MMORPG genre emphases in social interactions. As final remarks, from the total of 27 characteristics of the six human needs, as depicted by (BOSTAN, 2009), 19 were used in the action-need map; the game content does not approach the Sensual need; some actions attain or impair more than one human need; and finally, from the total of 87 preprocessed actions, only 50 were linked to human needs (obeying the aforementioned special considerations).

Table 67 – Summary of attainments and impairments per human need according to the action-need map

Human Need	**Number of Attainment Actions**	**Number of Impairment Actions**
Materialism	8	3
Power	10	10
Affiliation	20	7
Achievement	14	5
Information	6	5
Sensual	0	0

It is expected that some actions have an occurrence rate greater than others. Thus, to highlight this data aspect, the frequency of each action to each subset was computed and ranked, being presented in Table 68 (the actions that are not linked to any human need have the "*" sign).

On the one hand, it is possible to notice that the top 10 most-performed actions for all subsets regard the same group of actions, with little difference in their rank, and on the other hand, there are some actions with few occurrences (a percentage near zero). Moreover, the Train subset does not have the action "FeedingResult", which was added to the game content in a posterior upgrade. Figure 41 summarizes the attainments and impairments of each human need, considering the total number of actions for each subset.

As we can see, there is little difference between the percentages of each human need group considering the three subsets, being the biggest one, a value of 5% (between Test_1 and Test_2 subsets regarding the Information need). Our understanding of this stable behavior is based on two linked rationales, where: (1) the game content offered the same degree of challenge to players over time, which entailed a similar consumption behavior;

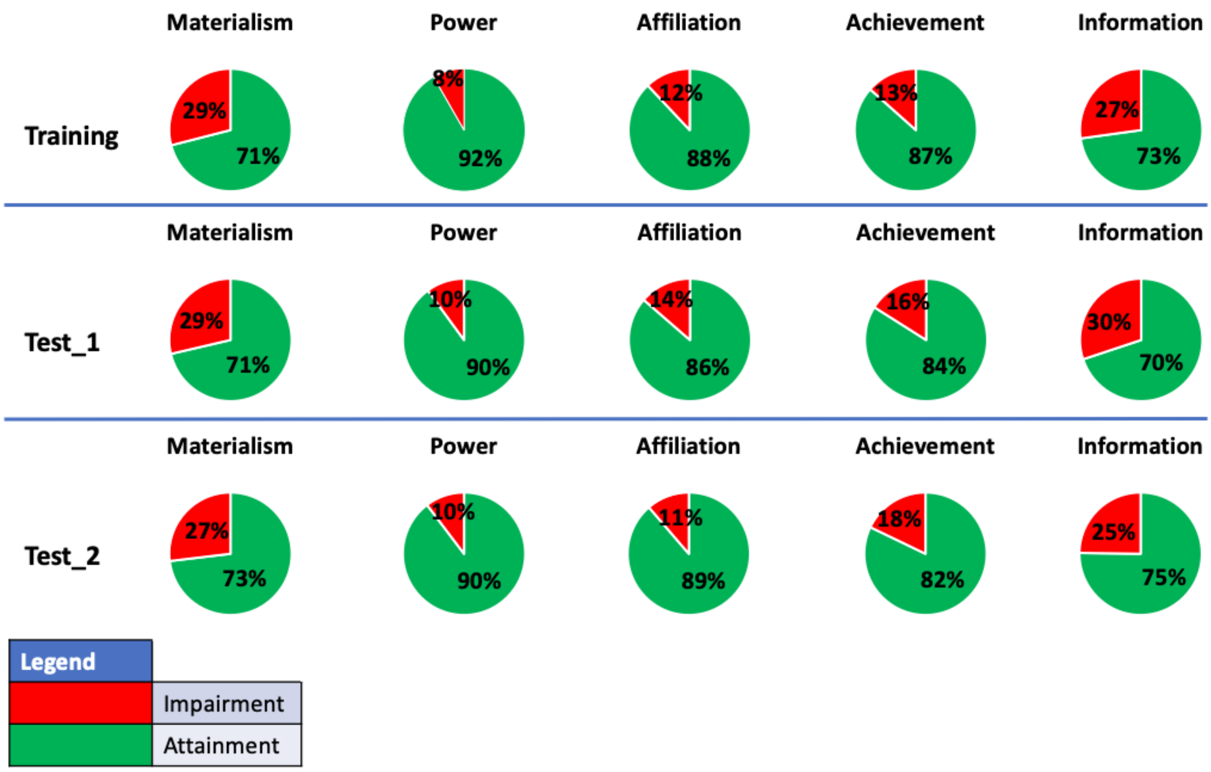


Figure 41 – Summary of attainments and impairments of human needs for each subset

Table 68 – Blade&Soul actions frequencies

Training		Test_1		Test_2	
Action	**Number of Occurrences**	**Action**	**Number of Occurrences**	**Action**	**Number of Occurrences**
GetMoney	22,087,783 (12.61%)	GetMoney	24,548,724 (12.42%)	GetItem	27,423,024 (13.26%)
Get Experience*	20,558,606 (11.74%)	Get Experience*	23,857,004 (12.07%)	Get Experience*	23,554,025 (11.39%)
KillNPC	20,343,746 (11.62%)	KillNPC	21,463,659 (10.86%)	GetMoney	23,509,201 (11.37%)
GetItem	19,115,057 (10.91%)	GetLoot Money*	20,603,786 (10.42%)	KillNPC	19,888,050 (9.62%)

To be continued

Training		**Test_1**		**Test_2**	
Action	**Number of Occurrences**	**Action**	**Number of Occurrences**	**Action**	**Number of Occurrences**
GetLoot Money*	18,189,480 (10.39%)	GetItem	20,321,443 (10.28%)	GetLoot Money*	18,313,001 (8.86%)
LoseItem	13,266,356 (7.57%)	LoseItem	14,156,970 (7.16%)	LoseItem	17,069,771 (8.26%)
GetItemBy Decomposition*	7,765,992 (4.43%)	GetItemBy Decomposition*	8,261,373 (4.18%)	GetItemBy Decomposition*	13,136,614 (6.35%)
LeaveZone*	5,105,848 (2.92%)	LeaveZone*	6,452,146 (3.26%)	LeaveZone*	6,850,166 (3.31%)
EnterZone*	5,073,545 (2.9%)	EnterZone*	6,414,143 (3.25%)	EnterZone*	6,809,232 (3.29%)
SpendMoney	4,166,112 (2.38%)	SpendMoney	4,507,274 (2.28%)	GetQuest Item*	4,240,403 (2.05%)
AcquireQuest*	3,416,005 (1.95%)	AcquireQuest*	4,060,923 (2.05%)	AcquireQuest*	3,995,098 (1.93%)
UseItem*	3,240,182 (1.85%)	UseItem*	3,367,431 (1.7%)	UseItem*	3,545,710 (1.71%)
GetQuest Item*	3,024,428 (1.73%)	GetQuest Item*	3,325,631 (1.68%)	Complete Quest	3,092,550 (1.5%)
Complete Quest	2,467,632 (1.41%)	SkillLevelUp	3,302,019 (1.67%)	Distribute Auction Money	2,777,291 (1.34%)
SkillLevelUp	2,316,085 (1.32%)	LearnTraining	2,937,934 (1.49%)	DepositItem	2,760,222 (1.33%)
LearnTraining	2,199,840 (1.26%)	Complete Quest	2,785,525 (1.41%)	SkillLevelUp	2,349,781 (1.14%)
LootItem*	2,134,508 (1.22%)	LeaveWorld*	2,046,337 (1.04%)	SpendMoney	2,270,806 (1.1%)
Distribute Auction Money	1,750,356 (1.%)	SaveEquip Info*	1,937,161 (.98%)	SaveEquip Info*	1,986,955 (.96%)

To be continued

Training		Test_1		Test_2	
Action	**Number of Occurrences**	**Action**	**Number of Occurrences**	**Action**	**Number of Occurrences**
SaveEquip Info*	1,675,045 (.96%)	Distribute Auction Money	1,836,597 (.93%)	LootItem*	1,937,306 (.94%)
DepositItem	1,535,765 (.88%)	LootItem*	1,810,253 (.92%)	LeaveWorld*	1,856,025 (.9%)
LeaveWorld*	1,496,071 (.85%)	DepositItem	1,776,729 (.9%)	AcquireSkill	1,311,060 (.63%)
EnterWorld*	1,043,490 (.6%)	Ressurrect*	1,451,176 (.73%)	Ressurrect*	1,299,457 (.63%)
Ressurrect*	987,096 (.56%)	EnterWorld*	1,367,183 (.69%)	UnEquipItem*	1,275,444 (.62%)
UnEquip Item*	965,377 (.55%)	Die	1,259,409 (.64%)	EnterWorld*	1,259,451 (.61%)
Dispose Quest	794,157 (.45%)	UnEquip Item*	1,253,996 (.63%)	Die	1,047,713 (.51%)
Die	765,050 (.44%)	Dispose Quest	1,169,450 (.59%)	BidParty Auction*	941,784 (.46%)
SellItem*	677,057 (.39%)	JoinParty	904,841 (.46%)	SellItem*	896,369 (.43%)
Exhaustion	645,904 (.37%)	KillPC	813,720 (.41%)	JoinParty	827,194 (.4%)
JoinParty	628,210 (.36%)	MoveTo Arena*	689,751 (.35%)	Dispose Quest	805,599 (.39%)
PartyAuction Start*	614,515 (.35%)	SellItem*	664,036 (.34%)	RetriveItem	753,702 (.36%)
BidParty Auction*	570,801 (.33%)	PartyAuction Start*	643,591 (.33%)	PartyAuction Start*	729,017 (.35%)
MoveTo Arena*	459,179 (.26%)	Exhaustion	641,097 (.32%)	Exhaustion	723,259 (.35%)
Teleport*	439,826 (.25%)	BidParty Auction*	595,770 (.3%)	LearnTraining	687,641 (.33%)

To be continued

Training		**Test_1**		**Test_2**	
Action	**Number of Occurrences**	**Action**	**Number of Occurrences**	**Action**	**Number of Occurrences**
KillPC	432,861 (.25%)	RetriveItem	518,950 (.26%)	MoveTo Arena*	618,214 (.3%)
RetriveItem	430,692 (.25%)	DismissParty	508,252 (.26%)	Teleport*	606,961 (.29%)
DismissParty	429,503 (.25%)	Teleport*	492,680 (.25%)	DismissParty	525,547 (.25%)
PartyAuction Success*	413,349 (.24%)	PartyAuction Success*	464,027 (.23%)	PartyAuction Success*	505,169 (.24%)
ResultOf Transform	390,973 (.22%)	ResultOf Transform	374,862 (.19%)	KillPC	482,996 (.23%)
PutMain Auction*	332,280 (.19%)	GetItem FromNPC	373,420 (.19%)	ResultOf Transform	457,229 (.22%)
BuyItemNow MainAuction*	319,122 (.18%)	GrowUpItem	338,898 (.17%)	GetItem FromNPC	401,537 (.19%)
GrowUpItem	313,852 (.18%)	PutMain Auction*	320,087 (.16%)	UseGathering Item*	339,660 (.16%)
DestroyItem*	278,813 (.16%)	AcquireSkill	316,785 (.16%)	DestroyItem*	329,395 (.16%)
GetItem FromNPC	269,814 (.15%)	BuyItemNow MainAuction*	315,136 (.16%)	GetGathering Item*	307,992 (.15%)
GetGathering Item*	260,280 (.15%)	RepairItem	285,620 (.14%)	GetChallenge TodayItem*	268,725 (.13%)
UseGathering Item*	239,891 (.14%)	DestroyItem*	260,798 (.13%)	PutMain Auction*	256,146 (.12%)
RepairItem	227,720 (.13%)	GetGathering Item*	223,216 (.11%)	RepairItem	251,381 (.12%)
GetChallenge TodayItem*	194,082 (.11%)	InviteParty	215,479 (.11%)	BuyItemNow MainAuction*	240,393 (.12%)

To be continued

Training		**Test_1**		**Test_2**	
Action	**Number of Occurrences**	**Action**	**Number of Occurrences**	**Action**	**Number of Occurrences**
InviteParty	190,933 (.11%)	UseGathering Item*	212,373 (.11%)	InviteParty	208,490 (.1%)
DuelEnd(PC)	125,727 (.07%)	GetChallenge TodayItem*	204,017 (.1%)	GrowUpItem	180,391 (.09%)
AcquireSkill	121,618 (.07%)	PartyBattle Result(PC)	200,078 (.1%)	DuelEnd(PC)	159,310 (.08%)
Trade GiveItem	115,980 (.07%)	DuelEnd(PC)	176,004 (.09%)	PcLevelUp	96,487 (.05%)
OccupyBase	71,029 (.04%)	OccupyBase	141,227 (.07%)	PartyBattle Result(PC)	87,446 (.04%)
PcLevelUp	62,672 (.04%)	PcLevelUp	57,149 (.03%)	Complete Challenge Today	62,997 (.03%)
PartyBattle Result(PC)	58,033 (.03%)	Complete Challenge Today	53,136 (.03%)	OccupyBase	57,984 (.03%)
Complete Challenge Today	55,328 (.03%)	Trade GetItem	48,921 (.02%)	RefuseParty	40,661 (.02%)
Trade GetItem	50,515 (.03%)	Trade GiveItem	37,772 (.02%)	RevealItem	40,086 (.02%)
RevealItem	32,776 (.02%)	Trade Get-Money	37,333 (.02%)	ConsumeGem ByReveal*	40,086 (.02%)
ConsumeGem ByReveal*	32,776 (.02%)	Trade Give-Money	34,240 (.02%)	Feeding Result*	33,324 (.02%)
Trade Give-Money	29,910 (.02%)	RevealItem	33,952 (.02%)	Trade Give-Money	33,208 (.02%)
Trade Get-Money	27,470 (.02%)	ConsumeGem ByReveal*	33,952 (.02%)	Trade Get-Money	30,871 (.01%)

To be continued

Training		**Test_1**		**Test_2**	
Action	**Number of Occurrences**	**Action**	**Number of Occurrences**	**Action**	**Number of Occurrences**
GetQuest Skill	25,903 (.01%)	PartyBattle End(Team)	33,792 (.02%)	GetQuest Skill	28,381 (.01%)
RefuseParty	15,451 (.01%)	GetQuest Skill	23,941 (.01%)	Expand Warehouse	27,844 (.01%)
Expand Warehouse	13,447 (.01%)	RefuseParty	19,662 (.01%)	DuelEnd (Team)	21,149 (.01%)
DuelEnd (Team)	9,827 (.01%)	DuelEnd (Team)	19,284 (.01%)	Trade GiveItem	16,575 (.01%)
ExceedItem Limit*	9,559 (.01%)	Expand Warehouse	12,028 (.01%)	Trade GetItem	15,980 (.01%)
PartyBattle End(Team)	9,530(.01%)	ExceedItem Limit*	9,295 (≈0%)	PartyBattle End(Team)	14,838 (.01%)
KickParty Member	5,823 (≈0%)	Feeding Result*	9,095 (≈0%)	BuyMy Item*	7,387 (≈0%)
BuyMy Item*	5,109 (≈0%)	BuyMy Item*	6,148 (≈0%)	InviteGuild	6,256 (≈0%)
InviteGuild	4,259 (≈0%)	KickParty Member	5,985 (≈0%)	JoinGuild	6,178 (≈0%)
JoinGuild	3,849 (≈0%)	InviteGuild	4,207 (≈0%)	DeletePC*	6,156 (≈0%)
DeletePC*	3,394 (≈0%)	DeletePC*	3,630 (≈0%)	KickParty Member	6,042 (≈0%)
Dissmiss Guild	3,038 (≈0%)	JoinGuild	3,019 (≈0%)	ChangeItem Look	5,497 (≈0%)
ChangeItem Look	1,298 (≈0%)	Dissmiss Guild	2,159 (≈0%)	Dissmiss Guild	4,159 (≈0%)
KickGuild Member	902 (≈0%)	ChangeItem Look	1,561 (≈0%)	ExceedItem Limit*	3,663 (≈0%)
RefuseGuild Invite	566 (≈0%)	KickGuild Member	1,041 (≈0%)	KickGuild Member	1,474 (≈0%)

To be continued

Training		**Test_1**		**Test_2**	
Action	**Number of Occurrences**	**Action**	**Number of Occurrences**	**Action**	**Number of Occurrences**
ExpireEvent Item*	231 (≈0%)	ExpireEvent Item*	763 (≈0%)	RefuseGuild Invite	987 (≈0%)
Destroy Guild*	108 (≈0%)	RefuseGuild Invite	581 (≈0%)	Create Guild*	295 (≈0%)
Create Guild*	107 (≈0%)	Create Guild*	131 (≈0%)	Destroy Guild*	238 (≈0%)
Guild LevelUp	60 (≈0%)	Destroy Guild*	120 (≈0%)	Guild LevelUp	151 (≈0%)
-	-	Guild LevelUp	101 (≈0%)	ExpireEvent Item*	138 (≈0%)

and (2) there is a pattern of human needs chase that keeps the same over time[4]. In addition, Table 69 roughly describes general aspects for attaining or impairing each human need in the Blade&Soul game. It is interesting to highlight that, on the one hand, an impairment regarding Materialism, Affiliation, or Information demands the occurrence of a previous attainment (e.g., to spend money, it is needed to have money, or to dispose a quest or dismiss a party, first it is needed to acquire a quest or create a party respectively), while on the other hand, an impairment of Power or Achievement may happen without previous attainment. In conclusion, we consider the impairments over Power and Achievement riskier than the impairments over the other needs because it is possible to have situations where a player only impairs his/her needs during the gameplay, which may lead the player to abandon the game (see Section 3.4 for more details about risk situations).

Table 69 – General attainments and impairments of Blade&Soul

	Attainment	**Impairment**
Materialism	Item and money acquisition.	Item loss and money spending.
Power	The victory against NPCs or other players and the establishment of a social hierarchy.	The defeat against NPCs or other players and the dissolution of a social hierarchy.
Affiliation	Money sharing and party creation.	Party dismissal.
Achievement	Level up and learning/improvement of abilities.	Die.
Information	Quest completion and social interaction.	Quest disposing and party dismissal.

[4] We understand this second aspect as personality influence.

As a final part of the data preparation, the Commitment metric was computed for all players. As depicted in Section 3.7, this computation requires usage data with players' identification and the obtained score according to a time-span. After adapting the Blade&Soul data to the Commitment concepts, the following input vector V_i was proposed based on a weekly perspective.

$$\mathrm{V_i} = \{id_i, d_i, S_{i\_max\_normal}, S_{i\_max\_mastery}, lvlUp_{i\_qty}\}$$

Where id_i is the player $_i$ identification, d_i is the number of days played by player $_i$ in a week, $S_{i\_max\_normal}$ is the player $_i$ max normal level obtained in the week, $S_{i\_max\_mastery}$ is the player $_i$ max mastery level obtained in the week, and $lvlUp_{i\_qty}$ is the quantity of levels improved by player $_i$ during the week (which can be zero, but never negative).

As additional remarks, the Blade&Soul data depicts two kinds of levels for each player: the "normal level" and the "mastery level", where the mastery level only starts to grow when the avatar reaches the max normal level value. Moreover, there are situations where the same player can have more than one avatar; therefore, our input vector is slightly different from the one used by (KUMMER et al., 2016; KUMMER; NIEVOLA; PARAISO, 2017b; KUMMER; NIEVOLA; PARAISO, 2018b), as in their case, the data portrayed only one avatar per player. On their approach, the difference of the max level and the min level (Δs_i) was used to measure the number of levels that a player obtained, but as in the Blade&Soul case, the same player can have different avatars in different levels at the same time, the same computation may lead to an incorrect value; therefore we opted to represent a player's score evolution based on the number of times that he/she levels up any one of his/her avatars. Moreover, the $S_{i\_max\_normal}$ and $S_{i\_max\_mastery}$ consider only the avatars that a player used. We understand that players can create different avatars to try to enjoy the same game with another perspective (a new learning process over the game mechanisms), what may entertain them (entailing in a high commitment) or not. However, in our approach, if a player that has avatars in the max levels does not use them, they will not be considered during this player's commitment degree assessment, as we assume that the active avatars represent the current challenge to master the game mechanisms, being this challenge more relevant than the ones already achieved and abandoned.

Figure 42 illustrates the number of players on each commitment degree according to each week of each subset. Even though there are gaps of time between the subsets, the weeks are numbered as a sequence, where the range from 1 until 6 regards the Training, 7 until 15 the Test_1, and 16 until 23 the Test_2 subset. Moreover, the 15th-week presents a very distinct behavior compared to the others, as it has only one day of usage. Thus, to not add a bias in the overall analysis, we opted not to consider it.

It is possible to see that all weeks presented a number of high committed players

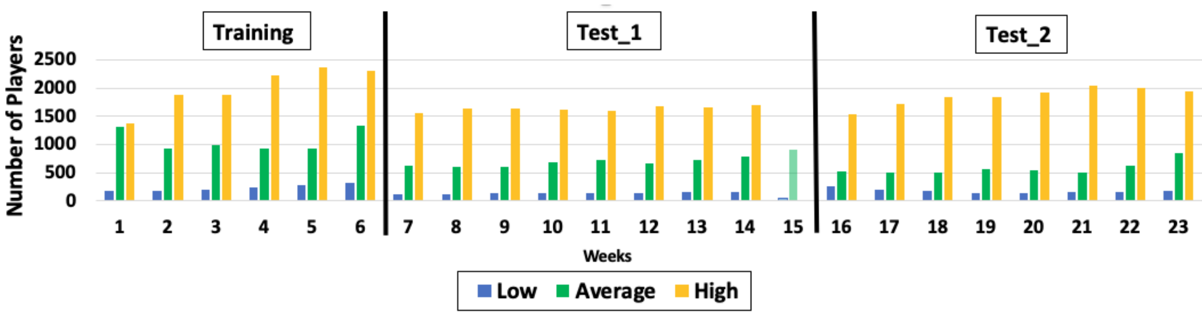


Figure 42 – Commitment assignment for Blade&Soul

greater than the other degrees, being the low committed one always with the least quantity. This disposition also highlights the "whales" characteristic of this dataset, where it is expected that the majority of players will be from the high commitment profile.

Moreover, by applying the Equation 5.6, the identified Maturity for the Blade&Soul dataset was 0.8, meaning that this dataset presents a low bias regarding the chase of human needs.

6.2 Technical Implementation

The original dataset was provided in a text file format. Firstly, all usage data was stored in a MySql database, and then preprocessing was performed. During this process, Stored Procedures and Java codes were developed and applied. The analyses of the data were performed through SQL queries.

After this initial analysis, the 3PIS was implemented. The 3PIS is a desktop application written in Java and has a strong linkage to MySql, where the processing is splitted between the Java code and Stored Procedures. All steps of the proposed method are parts of the 3PIS functionalities. In particular, we highlight an action-need map generator, which turns the only manual process of the proposed method simpler and straightforward. For more details, the system manual can be found in Appendix B.

6.3 Experimental Results

This section aims at presenting the results of the proposed method, which regards the psychological aspects of sentiments and personality traits (presented in Subsections 6.3.1, and 6.3.2, respectively). For each aspect, its results for the three subsets are presented, analyzed, and discussed. Also, individual players' behaviors are presented. All metrics computations considered a daily time-span perspective; thus, each metric has a maximum series size of 155 (the max number of days of the dataset).

6.3.1 Sentiments Results

Starting with the daily sentiment perspective, Figure 43 shows the daily sentiments of player "0021D8AA" of the Training subset. As we can see, this perspective allows the identification of a player's behavior in view of what is happening to this player, which can be focused on two main points: (1) what are the success or not of this player in attaining his/her needs, and (2) the chase or not of needs (an initial "clue" about this player's personality).

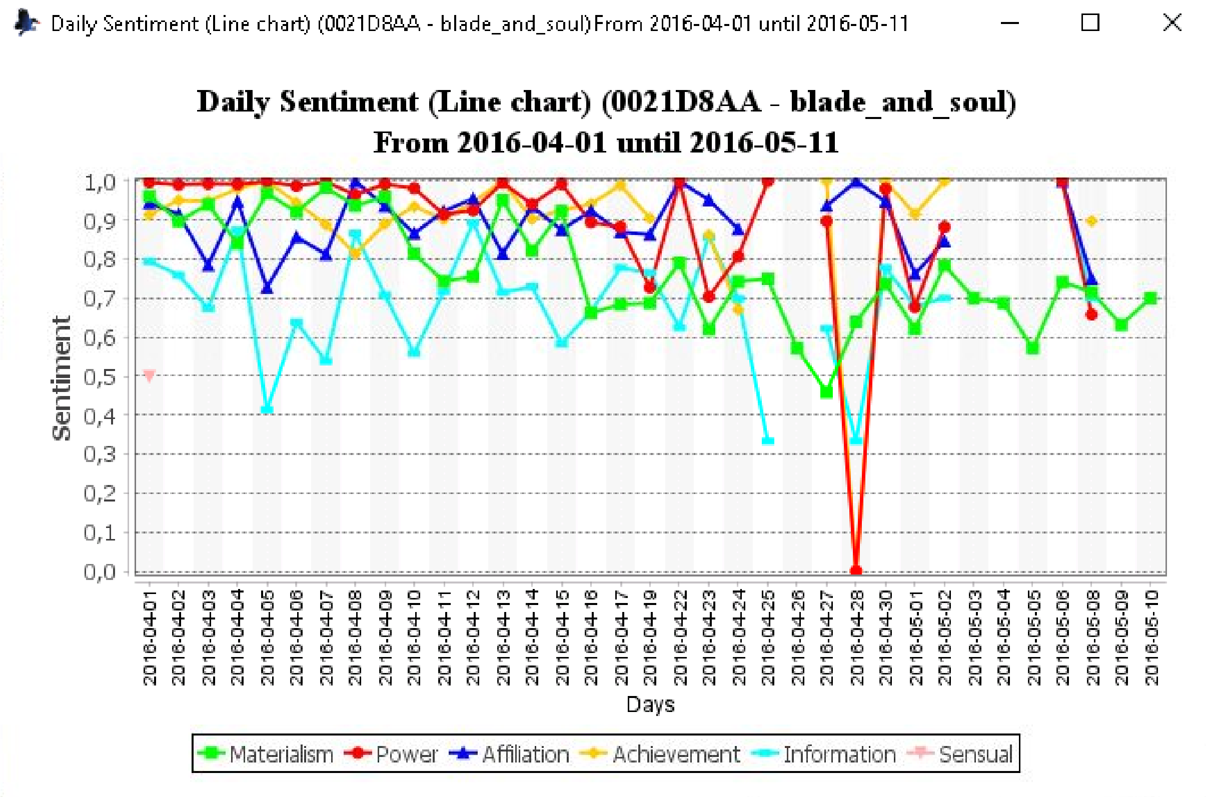


Figure 43 – Player 0021D8AA daily sentiment

On the one hand, it is possible to see that this player has positive sentiments (i.e., sentiment values above 0.5) in almost all days regarding all human needs, except by a few cases where the Materialism, Power, and Information needs presented negative sentiments, where the Power value of zero on day 2016-04-28 highlights that all power attempts on this day were not successful. On the other hand, we can see that the only need that is always chased is the Materialism one, as there are some days where only this need has sentiments linked to it. It shows this player's preference in chasing this kind of need. An additional remark regards the Sensual sentiment, as the game content does not provide the means to attains the Sensual need; by default, the players have a neutral sentiment on their first day regarding it.

Interestingly, there are cases where players do not chase the Materialism, Power, Affiliation, Achievement, and Information needs, such as shown by Figures 44, 45, 46, 47, and 48, regarding players 818BF259, 1C356FC0, B9BD0CF7, AE6D91F2, and 7BE2EBF2, respectively.

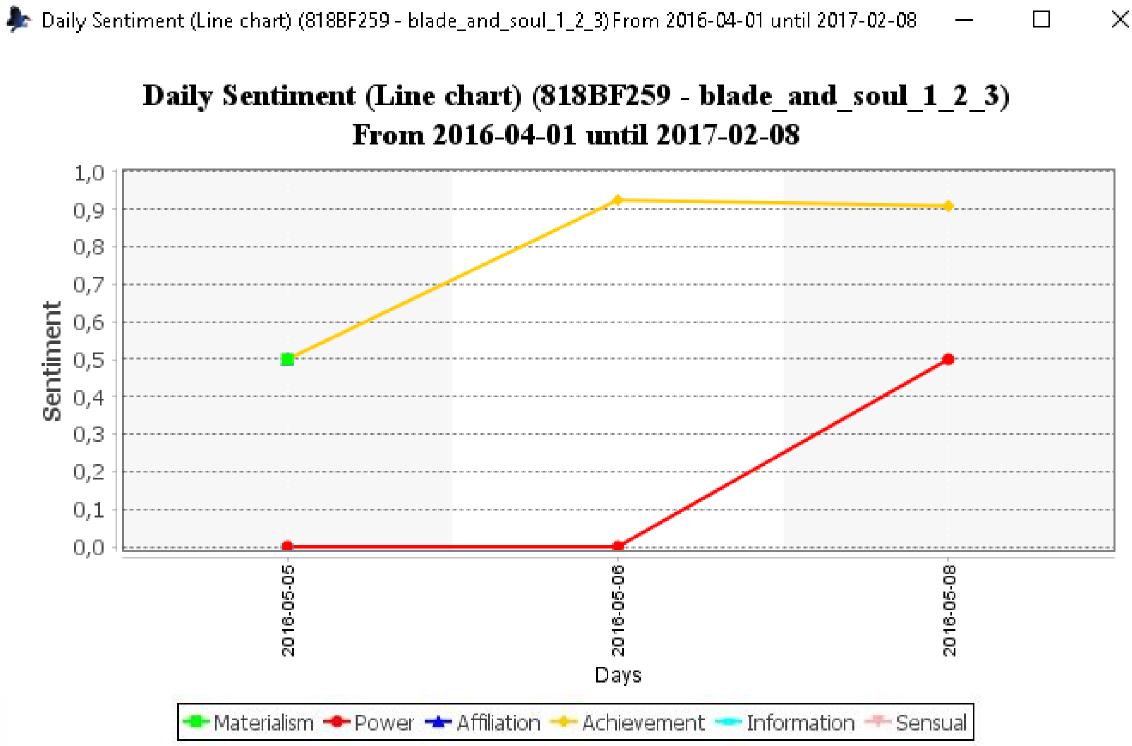


Figure 44 – Player 818BF259 daily sentiment - Materialism not chased

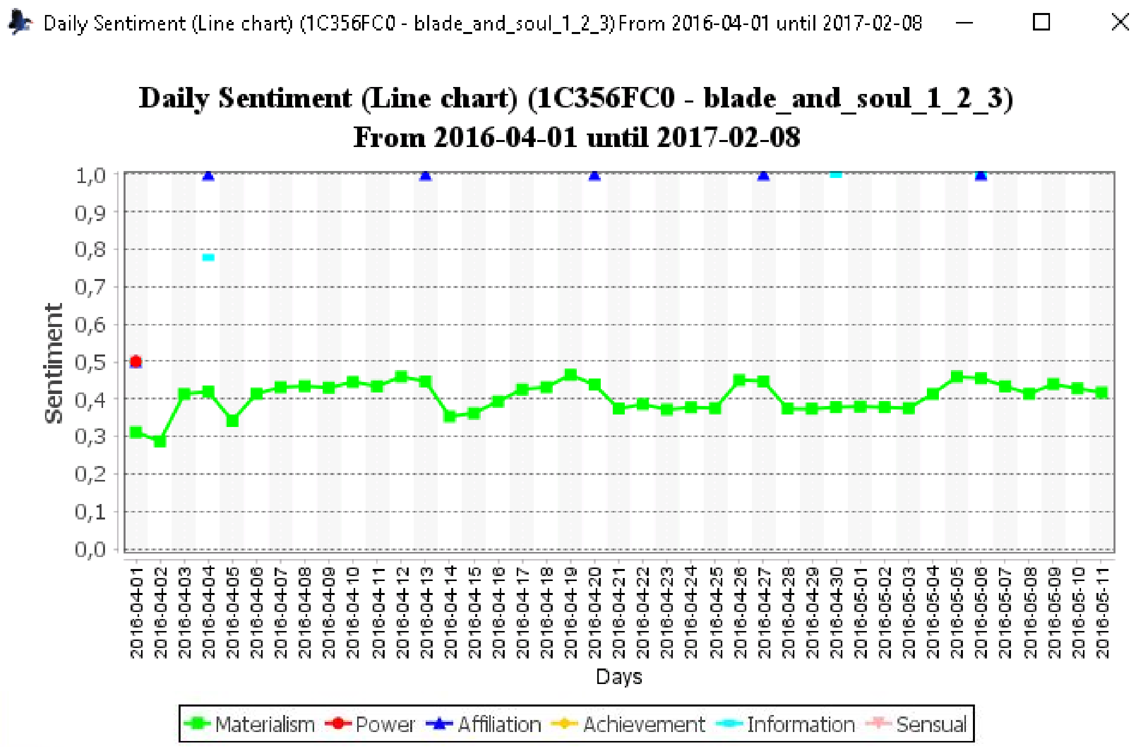


Figure 45 – Player 1C356FC0 daily sentiment - Power not chased

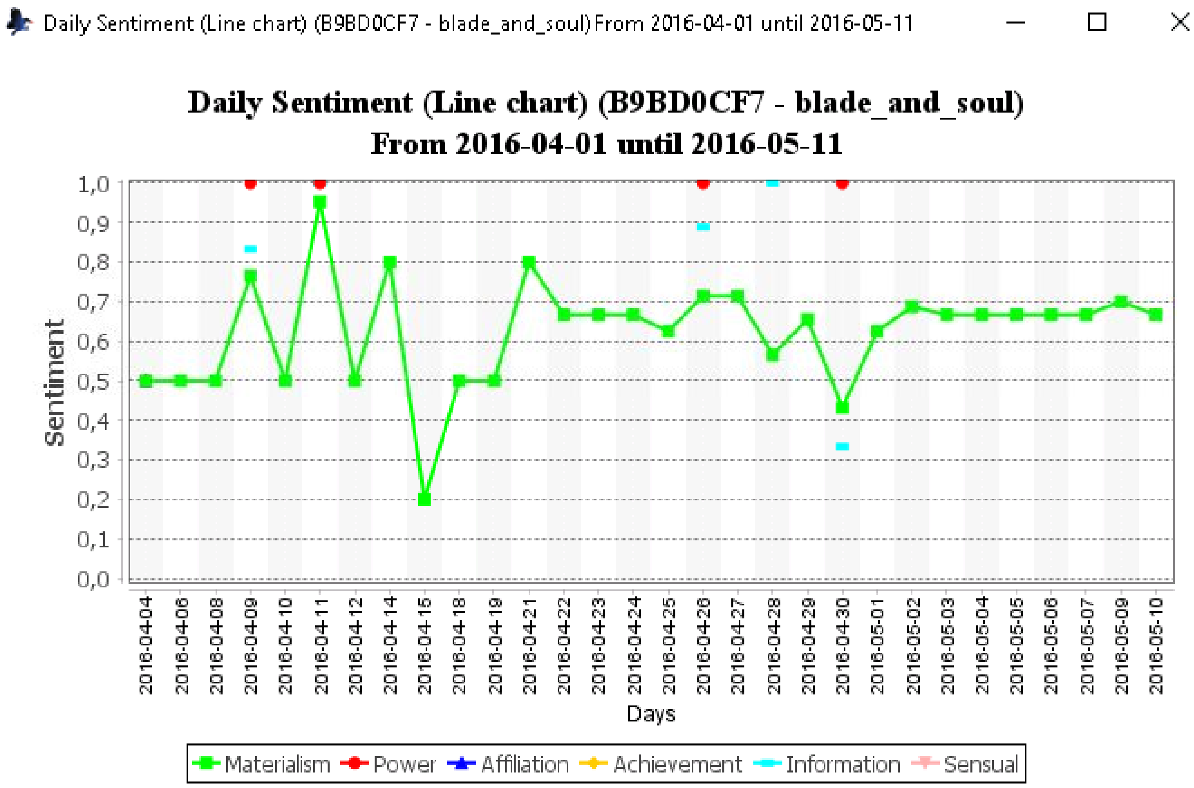


Figure 46 – Player B9BD0CF7 daily sentiment - Affiliation not chased

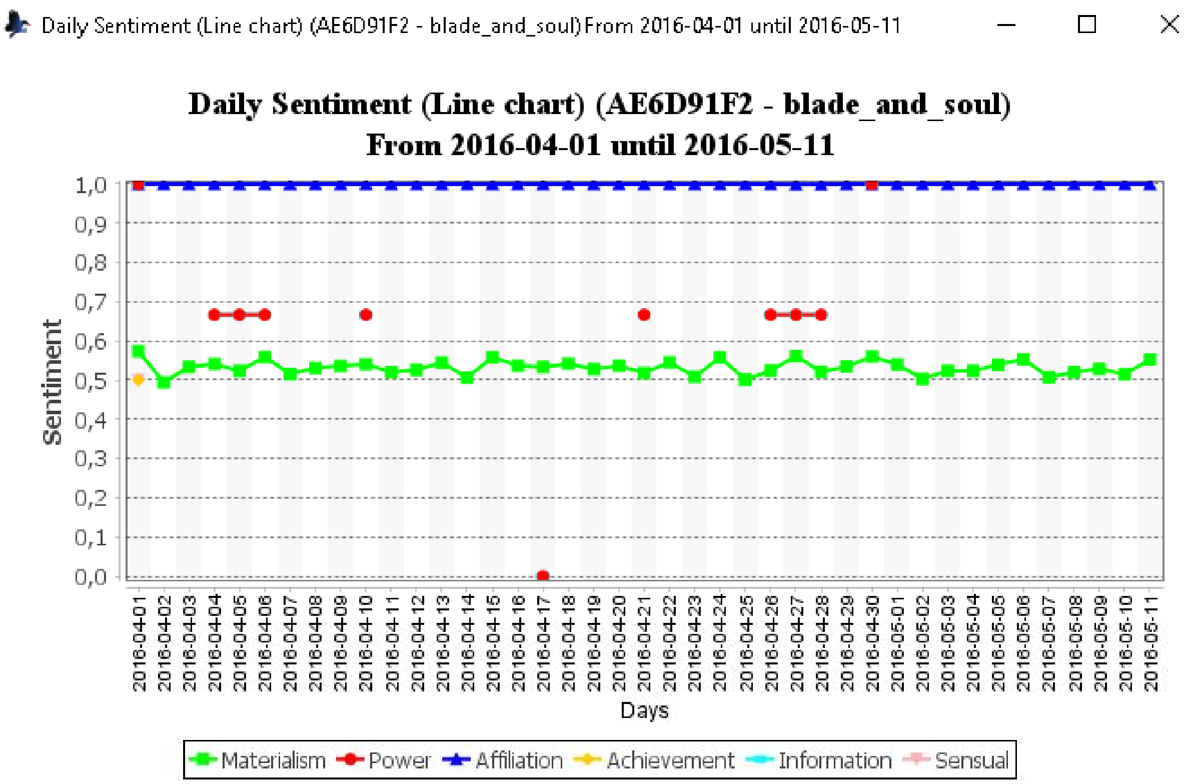


Figure 47 – Player AE6D91F2 daily sentiment - Achievement not chased

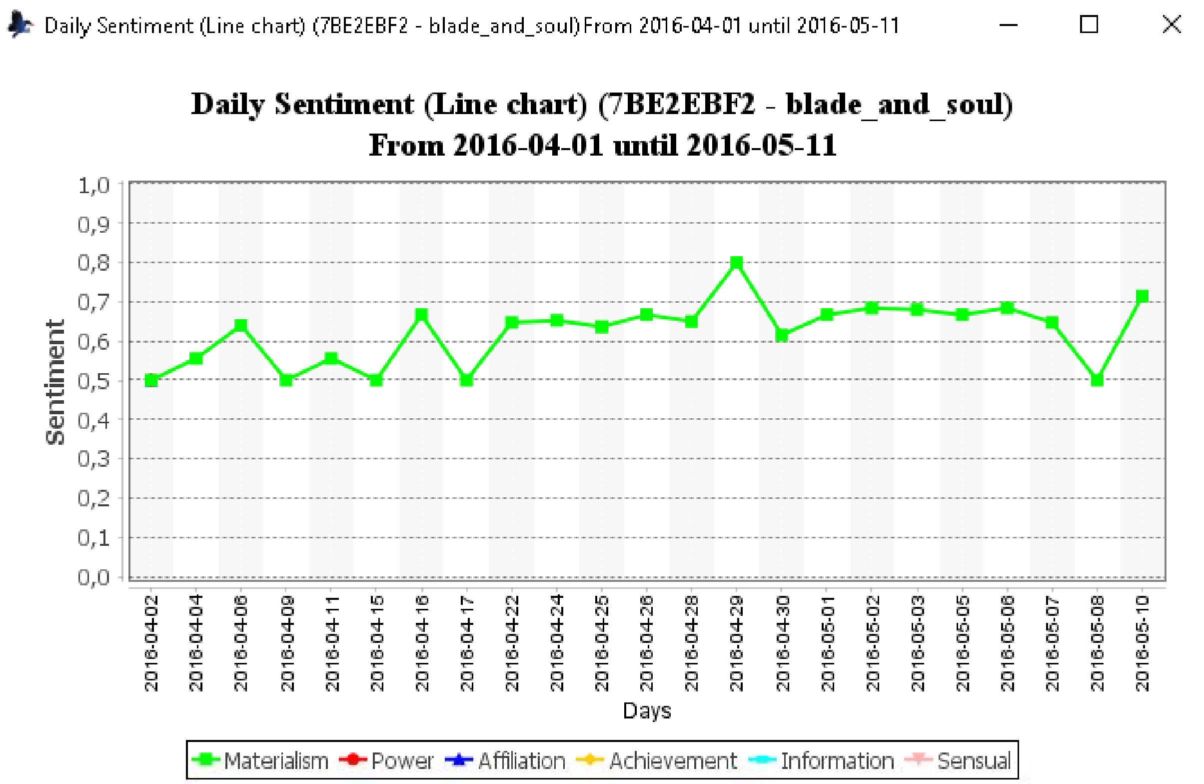


Figure 48 – Player 7BE2EBF2 daily sentiment - Information not chased

Attached to their different patterns of human needs chase, it is possible to see that they differ in the number of days played, and the number of human needs not chased (such as highlighted by players 818BF259, and 7BE2EBF2). Another point regards the sporadic chase of some needs, such as depicted by the sparse sentiment values of players 1C356FC0, B9BD0CF7, and AE6D91F2. An interesting point regards player AE6D91F2, where it is possible to see that he/she is a thriving social player. An overall view of all players daily sentiments from all subsets are presented in Figure 49. It is essential to highlight that a few players are present in more than one subset, such as depicted by Table 70.

Table 70 – Number of players present in more than one subset

Subsets			
1 and 2	**2 and 3**	**1 and 3**	**1, 2 and 3**
182	87	90	8

As we can notice, the most positive daily sentiment regards the Affiliation need, highlighting, once again, the importance that social aspects have in the MMORPG genre. The Sensual need presented daily sentiments on all days because for every day, there was at least one new player that carried the initial default neutral sentiment of this need. Moreover, the Achievement peaks refer to the moments where players improve their avatars, where the Test_1 peak refers to a game upgrade, and the Test_2 one to the change of the game business model from subscription to free-to-play, where a more significant amount

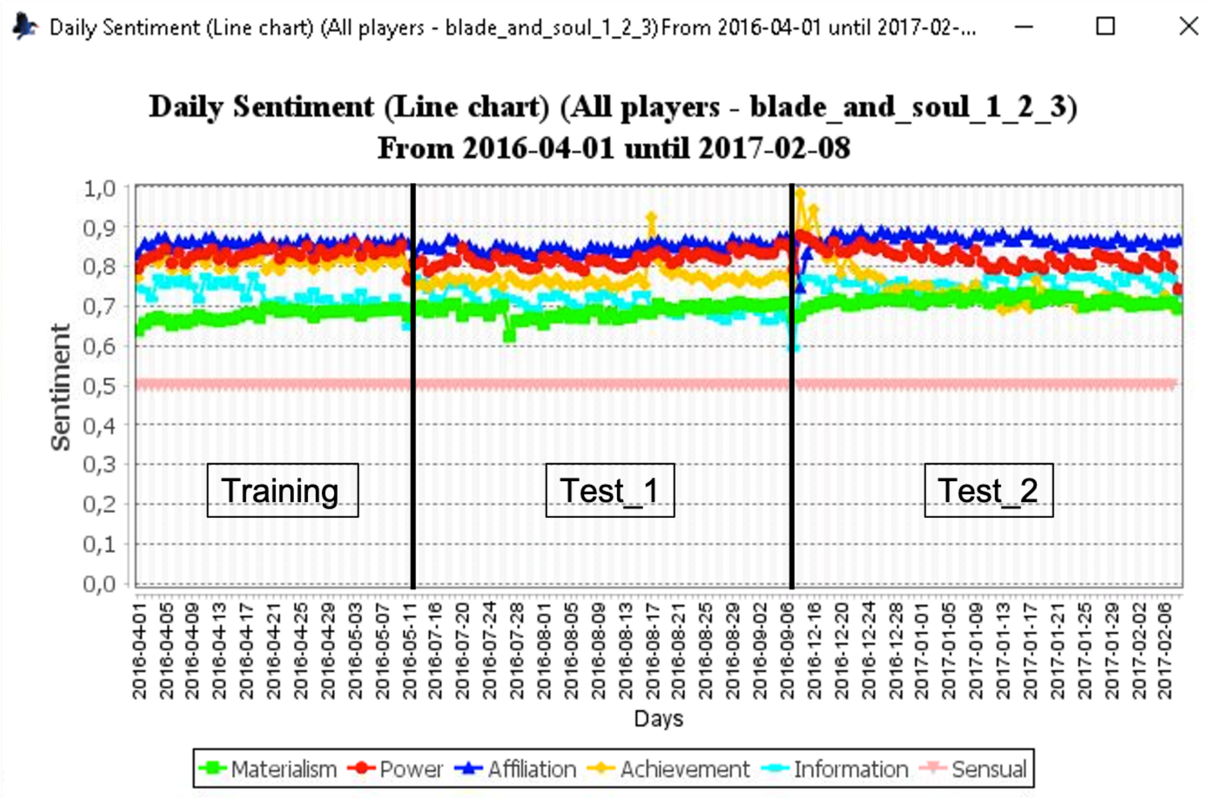


Figure 49 – All players daily sentiments for all subsets

of new players started to play at the same time (Table 71 presents the number of new players[5] per subset), entailing an increased degree of Achievement attainments. Another remark regards that, in general, all players' daily sentiments are positive. By linking to this fact, the "greener" color depicted by Figure 50 (a pie chart view, where the more positive the sentiments, the greener, whereas the more negative, the redder) and the abandonment[6] rate of Figure 51 [7], one possible conclusion is that this game has an acceptable balance between challenge degree and players sentiments, given that its players almost do not abandon the game and have, in general, a positive sentiment to the game.

Table 71 – New players per subset

Subset	# new players	# old players	Total of players
Train	1239	2761	4000
Test_1	1208	1792	3000
Test_2	1443	1557	3000

Moving to the historical sentiment perspective, Figures 52, 53, 54, 55, 56, and 57 portray the historical sentiments according to the same players depicted in the daily

[5] A player is labeled as new if he/she has played on level one.
[6] Abandonment in this context regards the last day when a player played.
[7] Remembering, this dataset does not contain all active players; thus, the abandonment rate at the end of each subset does not mean, precisely, that players are leaving the game; it just means that players are usually not present on more than one subset.

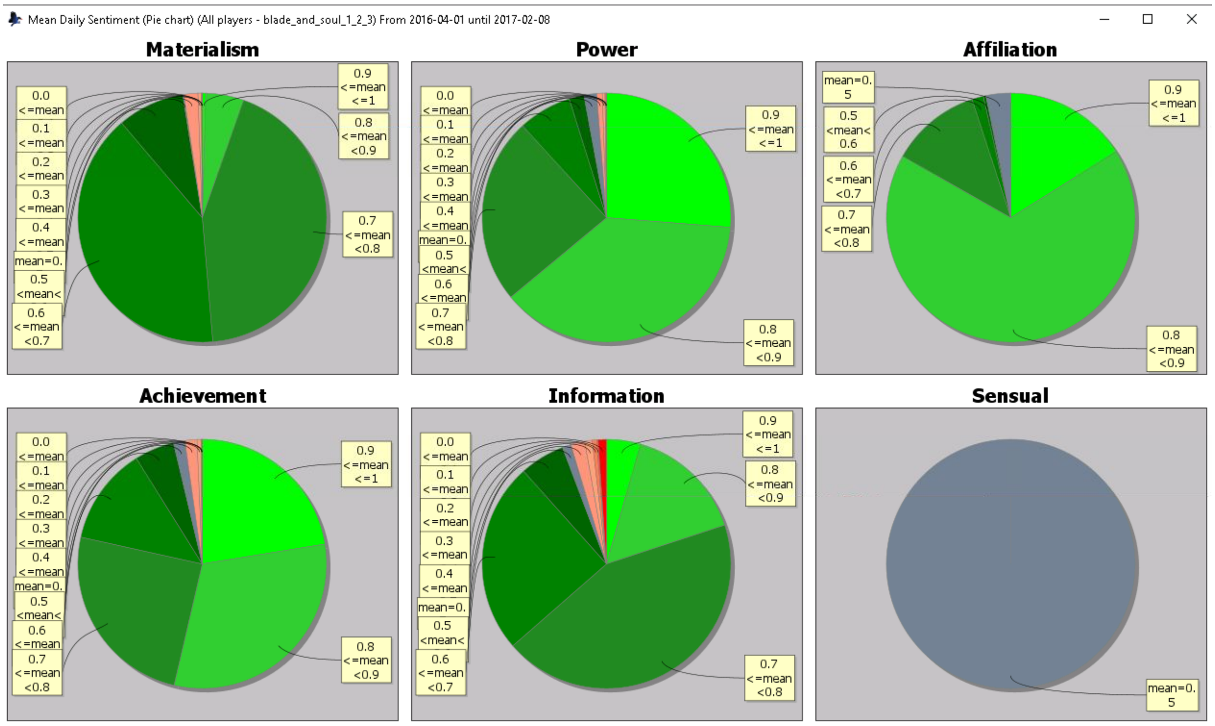


Figure 50 – Daily sentiments for all subsets - Pie chart

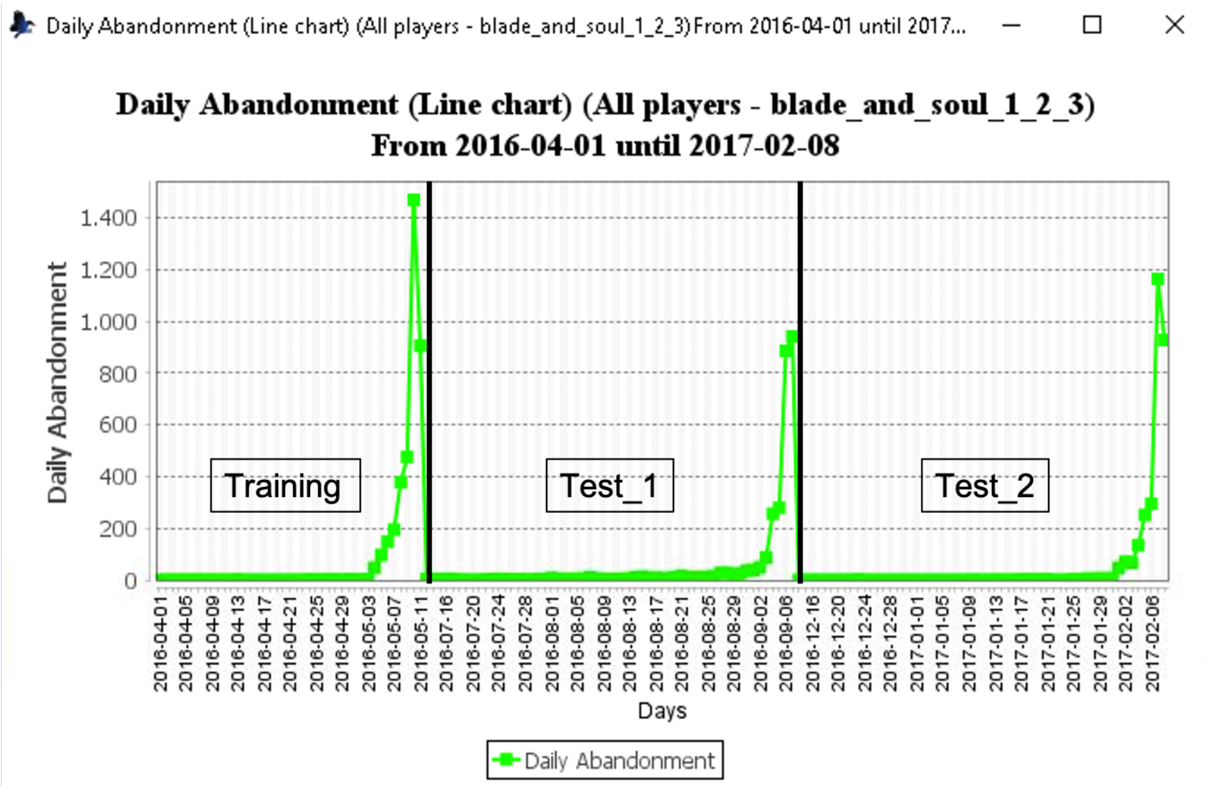


Figure 51 – Daily abandonment for all subsets

perspective, being they 0021D8AA, 818BF259, 1C356FC0, B9BD0CF7, AE6D91F2, and 7BE2EBF2, respectively.

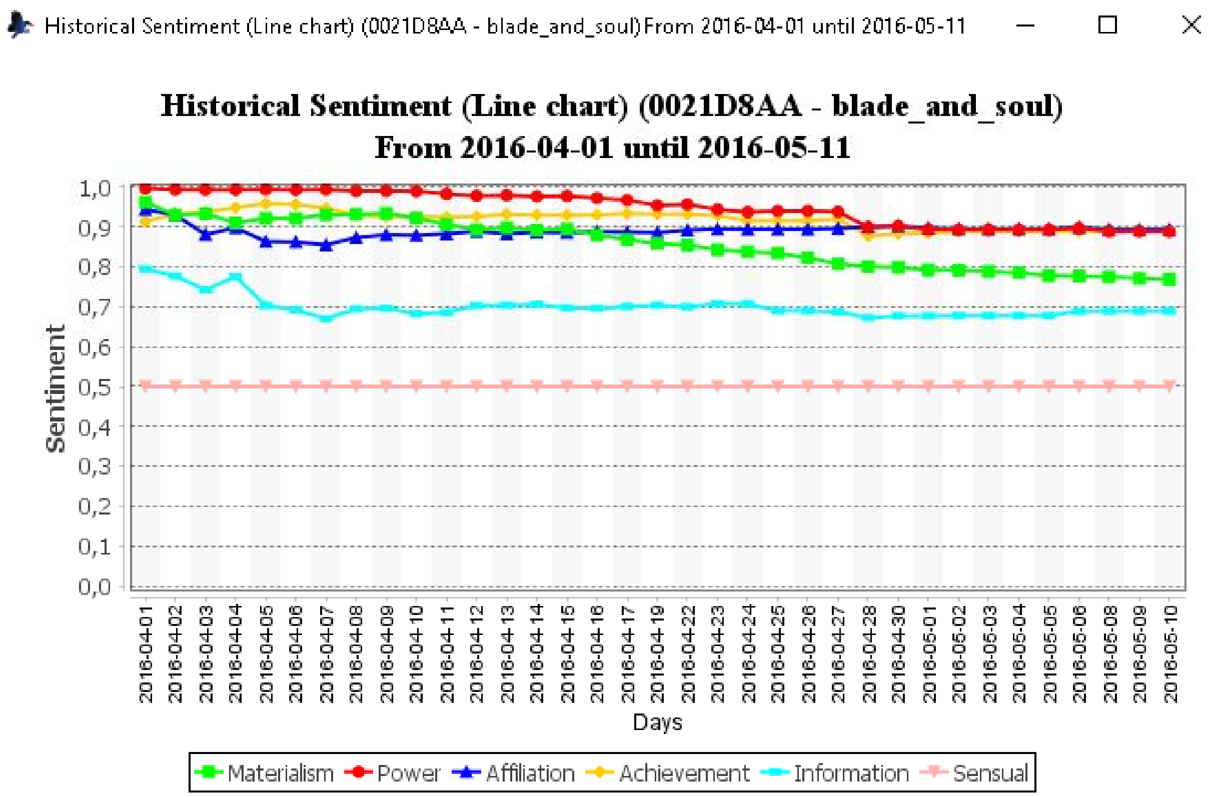


Figure 52 – Player 0021D8AA historic sentiment

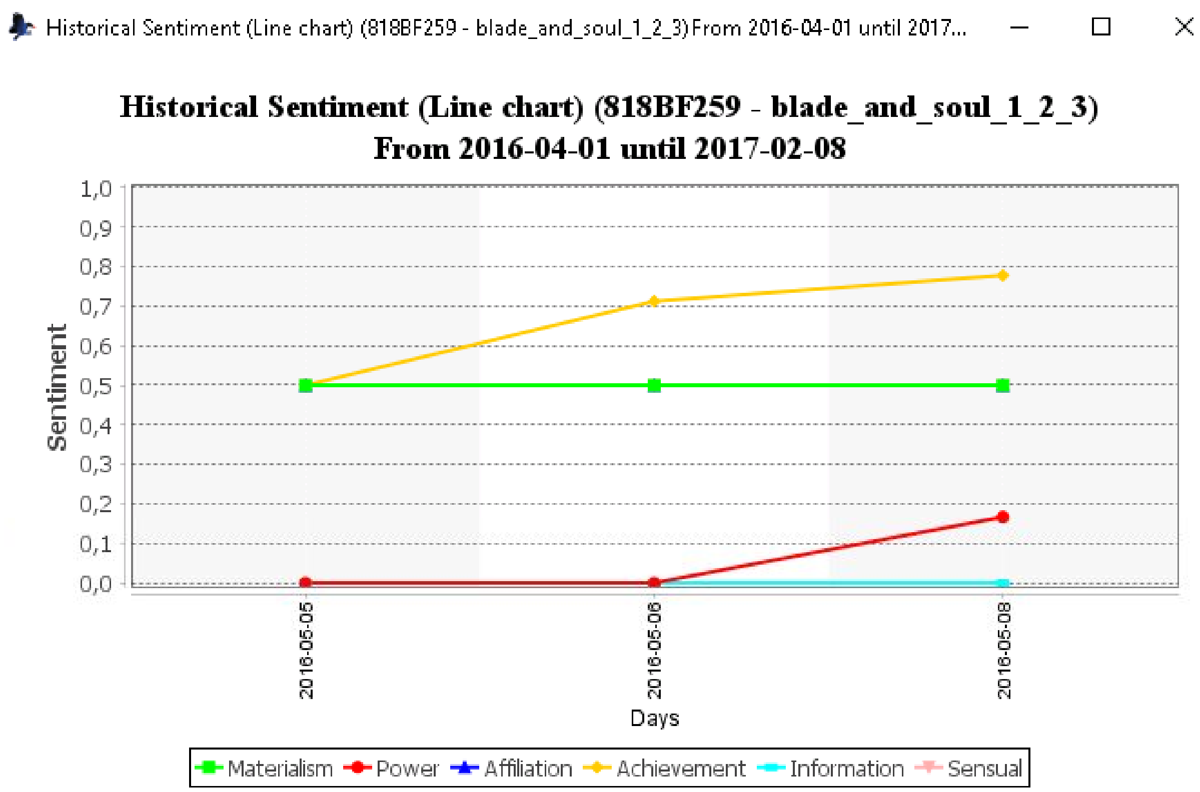


Figure 53 – Player 818BF259 historic sentiment - Materialism not chased

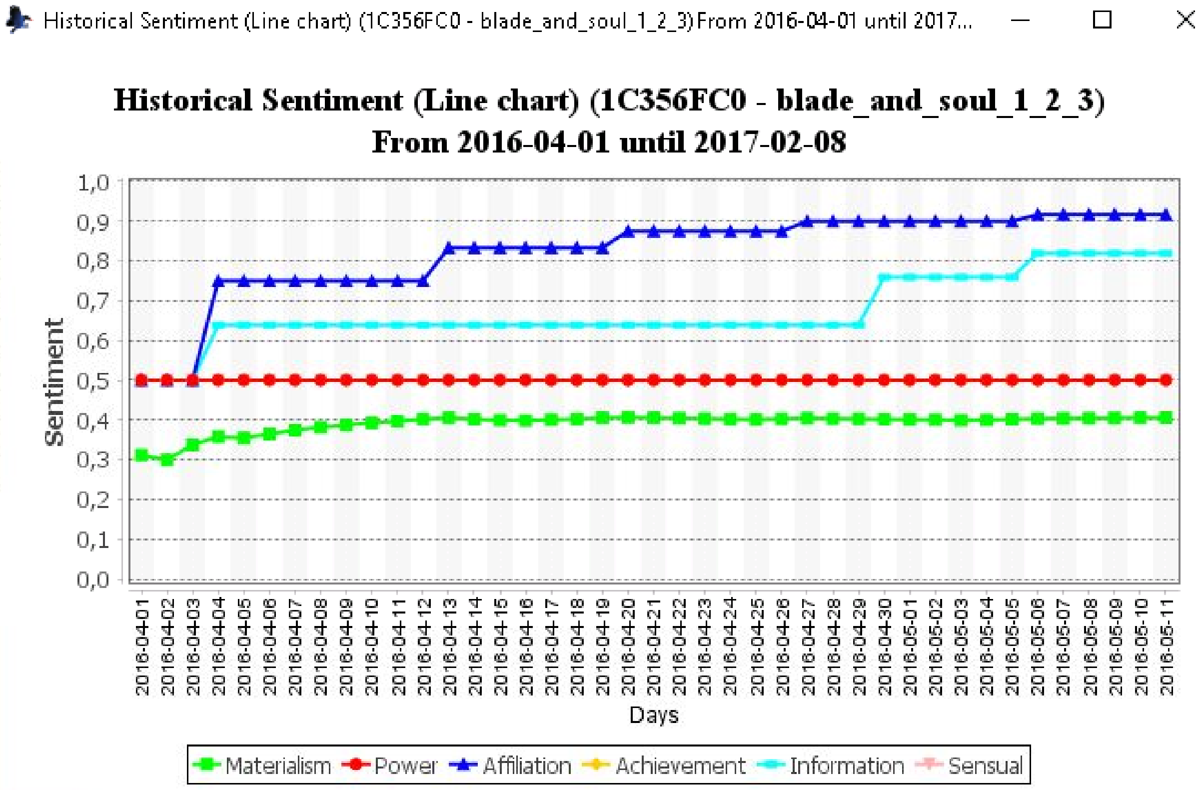


Figure 54 – Player 1C356FC0 historic sentiment - Power not chased

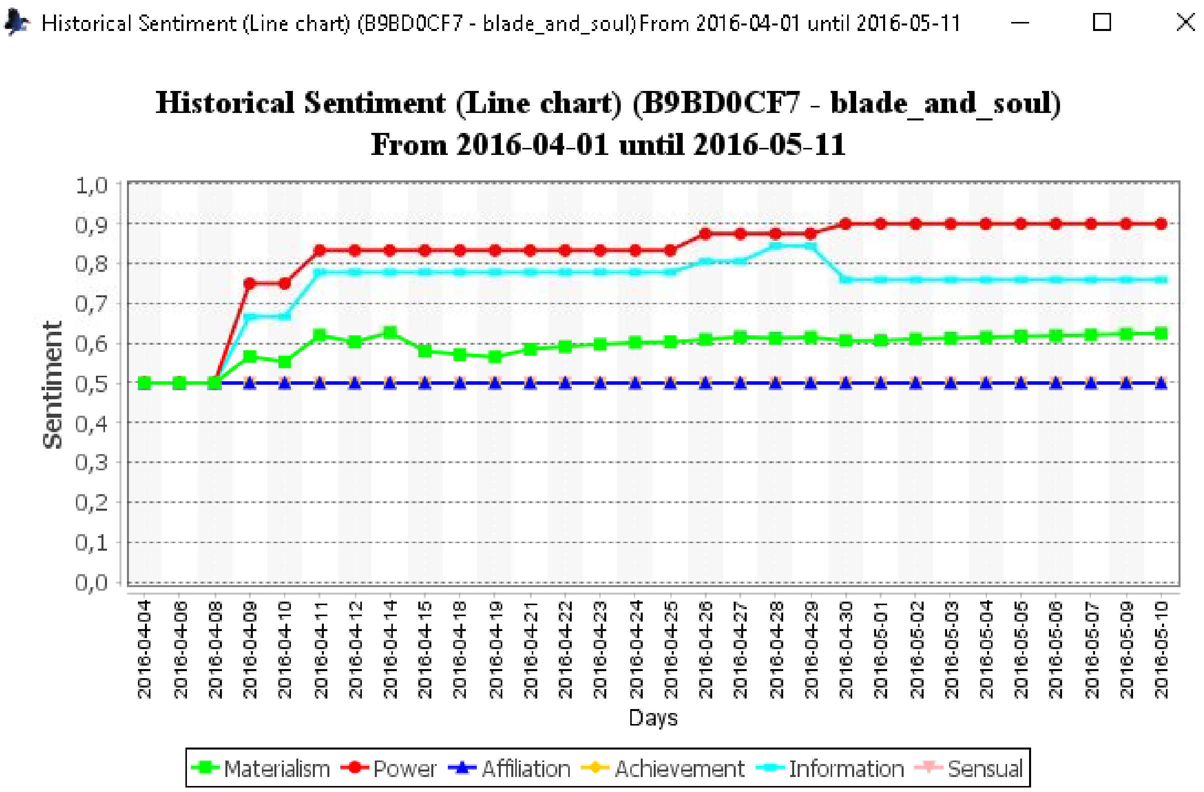


Figure 55 – Player B9BD0CF7 historic sentiment - Affiliation not chased

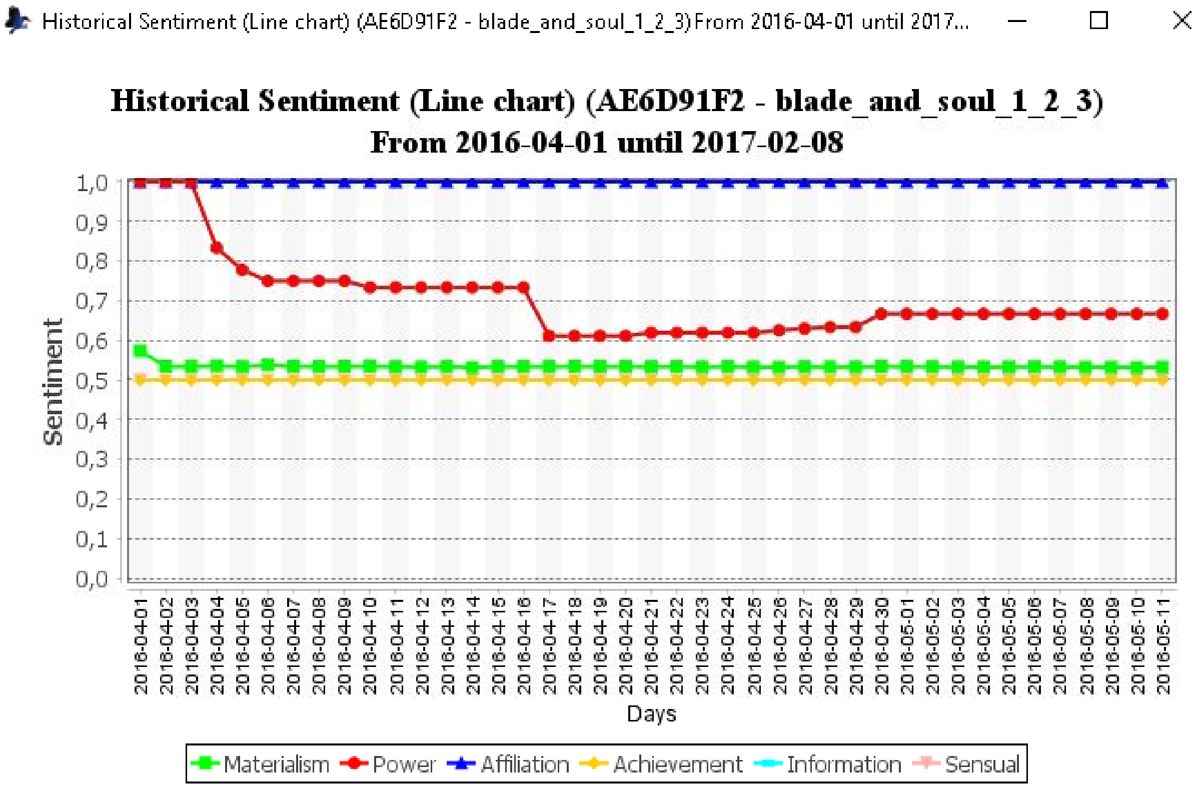


Figure 56 – Player AE6D91F2 historic sentiment - Achievement not chased

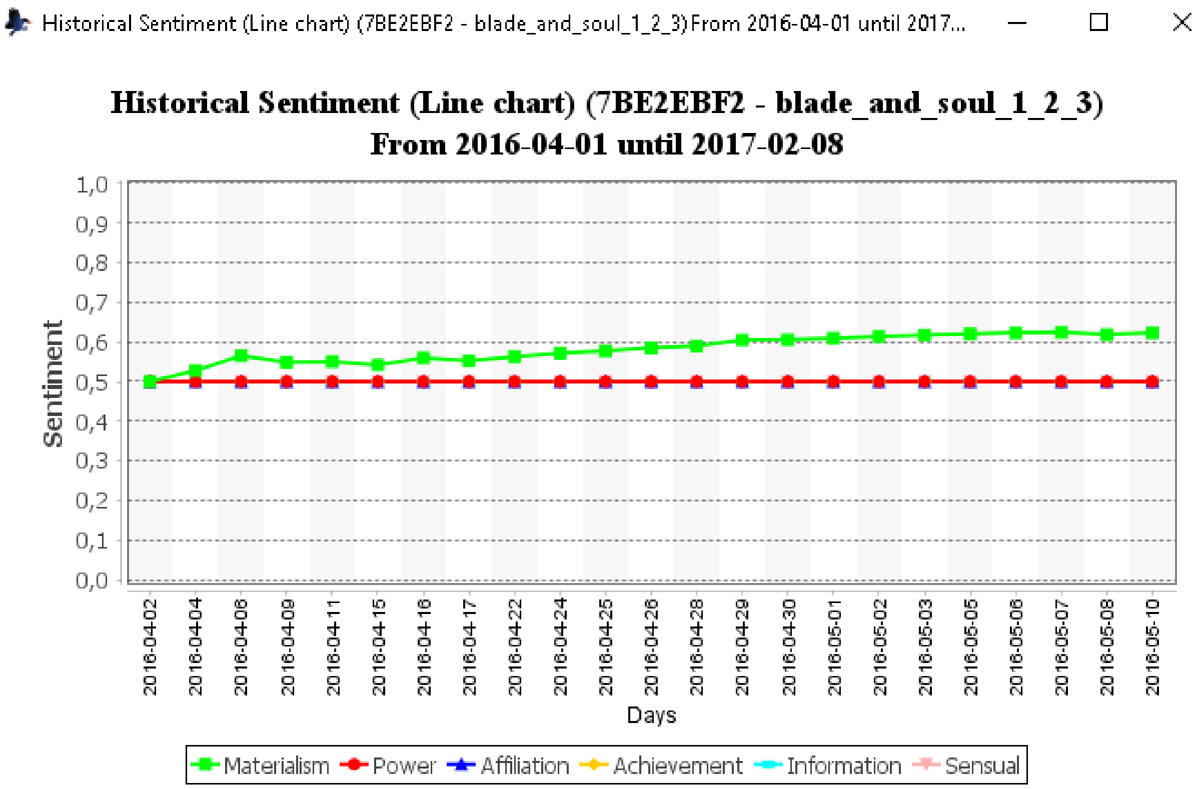


Figure 57 – Player 7BE2EBF2 historic sentiment - Information not chased

As we can see, in general, the players' sentiments are more stable in the historical perspective than on the daily one. Only the players 818BF259 and 1C356FC0 presented a negative value. Furthermore, all needs not chased are represented by a neutral sentiment (a value of 0.5). In particular, we highlight the player 0021D8AA that presented a daily sentiment of zero regarding the Power need on day 2016-04-28. In the historical perspective, his/her Power sentiment presented only a small decrease of 0.05, moving from 0.95 to 0.9. This fact shows the difference between the daily and historical perspectives, where even though a player can have a "bad day" regarding a given human need (i.e., a negative daily sentiment), in the overall context, this player can have a good sentiment to the same need. It is important to emphasize that analyzing these two perspectives together provides more details about players' behaviors, such as the persistence aspect. Assuming two players, A and B, where both A and B have positive historical sentiments regarding the Power need and a negative daily sentiment on the last day played regarding the same need, and A leaves the game whereas B stays. Given this, it is possible to state that B is more persistent than A.

As a final analysis of the players' sentiments, Figure 58 presents the historical sentiment considering all players from the three subsets. As only a few players are present in more than one subset, each subset's historical sentiments are mostly influenced by its own data. This fact is highlighted by the Achievement peak in Test_2 subset, which regards the new players' evolution (note that if a significant number of players was shared between all subsets, this peak would not be identified, at least, in the historical perspective). Another interesting fact about the Test_2 subset regards the Affiliation need, as it is possible to see a tendency of players chasing more this need according to their decrease in the Achievement need, meaning that players start to search for social interactions after having initial progress in-game.

6.3.2 Personality Traits Results

This Section is divided into three Subsections, regarding the Micro Spectrum (Subsection 6.3.2.1), the Macro Spectrum (Subsection 6.3.2.2), and the Micro and Macro Similarities (Subsection 6.3.2.3).

6.3.2.1 Micro Spectrum Results

The Micro Spectrum regards the Game Paths generation process, where according to the players' sequence of actions, each player is placed on a Game Path Segment. Such a generation carries a set of characteristics regarding splitting of segments, the players' personality influence (Equation 5.18), the number of Game Paths, the sharing of segments between different players, and the segments' lengths. Figures 59, 60, 61, 62, and 63 present all these characteristics regarding all subsets, respectively. It is essential to highlight that the Game Paths are generated following a chronology, where the Test_1 considers the

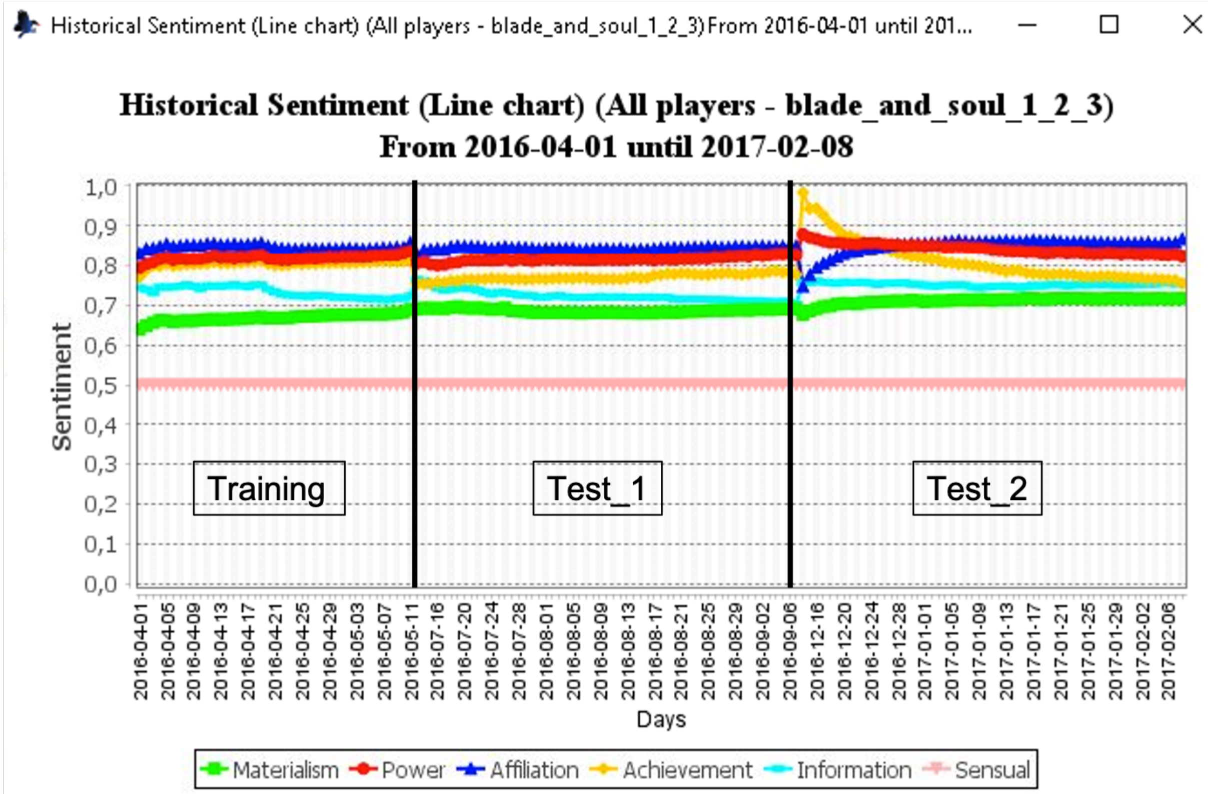


Figure 58 – All players historic sentiments for all subsets

Training Game Paths, as well as the Test_2 considers the Test_1 ones. It means that the Game Paths generated on the last day of Test_2 regards the full dataset behavior.

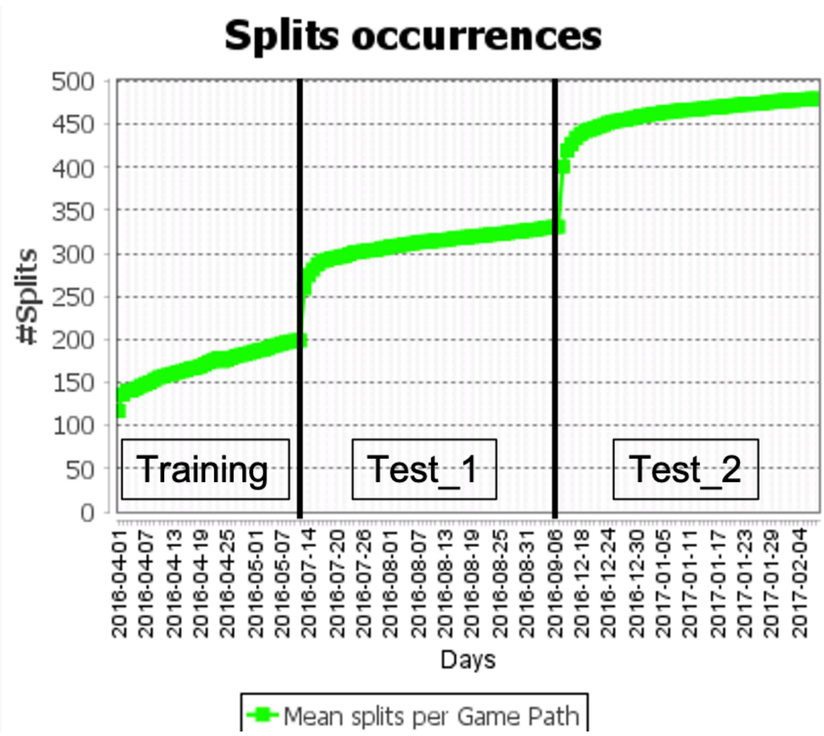


Figure 59 – Splits occurrences for all subsets

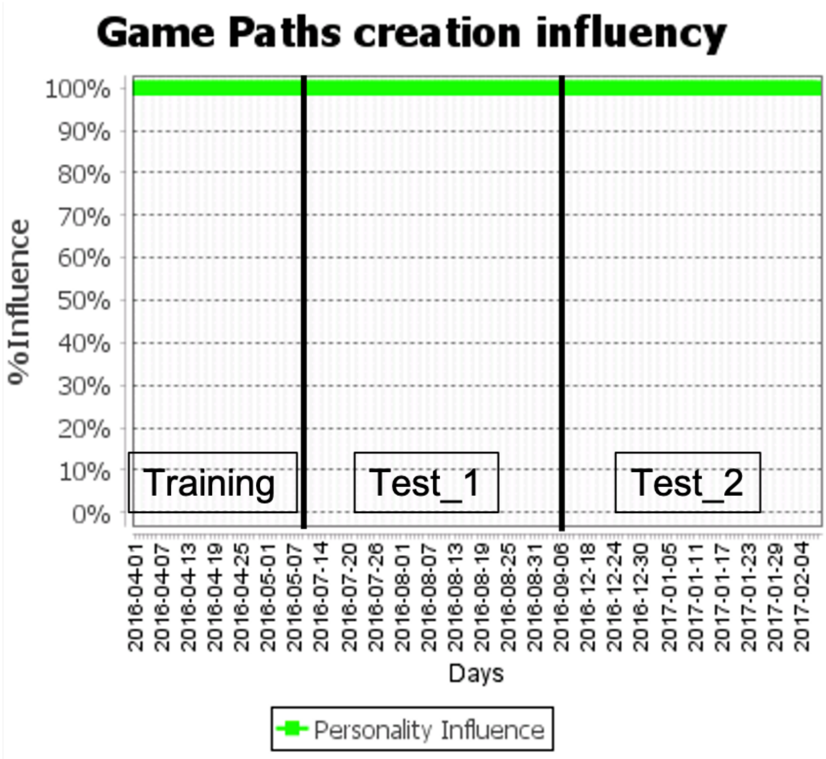


Figure 60 – General personality influence for all subsets

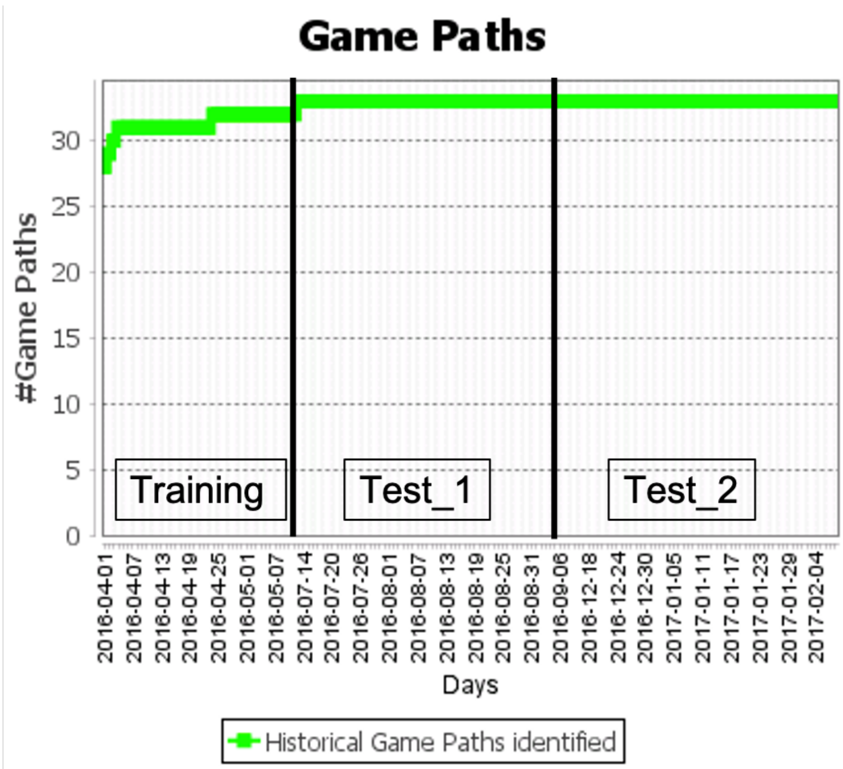
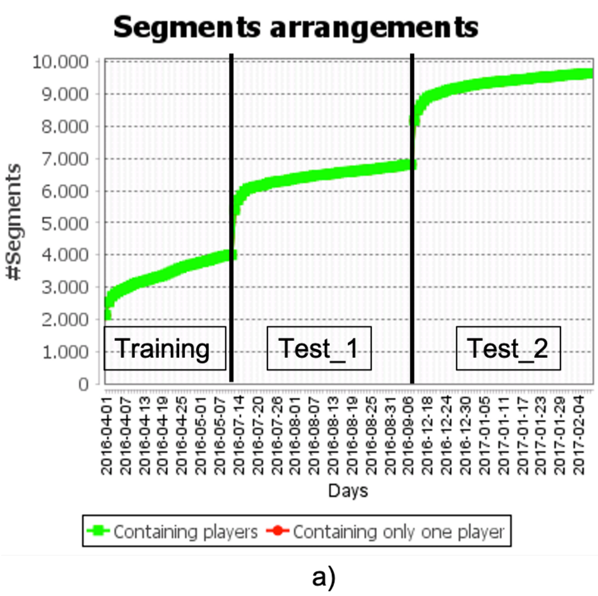


Figure 61 – Number of Game Paths for all subsets



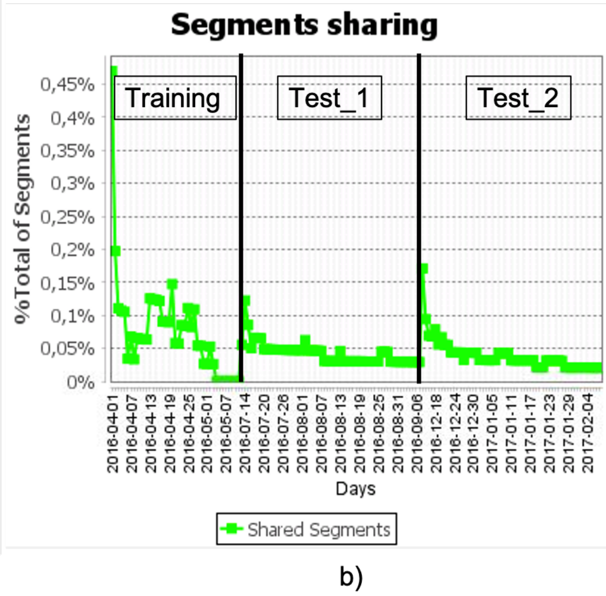


Figure 62 – a) Game Paths Segments arrangements; b) Game Paths Segments sharing

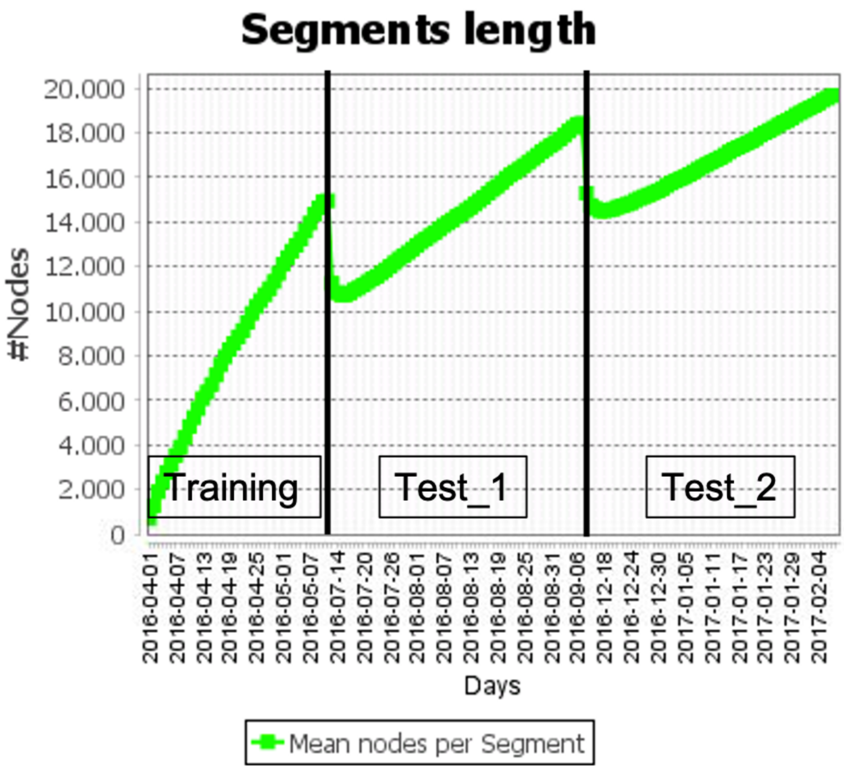


Figure 63 – The mean segments' length for all subsets

As we can see, Figure 59 shows a growth on the number of splits, highlighting the fact that each subset has distinct behaviors in the Micro Spectrum; Figure 60 presents a $GeneralPersInfluence$ of almost 100% for all subsets, which linked to the fact that there are more than one Game Path Segment, implies that players are playing based on their own desires, not being forced to follow one or another pattern of content consumption; Figure 61 exhibits an increment of the Game Paths generated in the Training according to the Test_1 subset behavior, where the number of Game Paths increased from 32 to 33, also, from the 50 possible actions, only 33 were chosen as the first action of a sequence; Figure 62 “a” part shows that almost all segments have only on player on it, meaning that, in the Micro Spectrum, almost all players are individually depicted, whereas the “b” part shows that the sharing of segments varied from 0.47% until 0%, fomenting the findings of “a” part, also, the increased values on the beginning of each subset (in the “b” part) highlights the fact that the players’ behaviors become more distinct in the pace that they play longer (what is expected); and Figure 63 portrays the variances of the mean segment length, where at the beginning of subsets Test_1 and Test_2 this value is decreased due to the fact that their players played for less time (i.e., have a shorter sequence of actions) than the players on each previous subset, also, the final value for subset Test_2 is similar to the one of Test_1, showing that the players from these subsets present a similar behavior in view of the amount of performed actions.

Another point of analysis of the Micro Spectrum regards the number of references each Position receives over time. As players can be placed on different segments over time, and each segment can vary in deepness (i.e., Position), each player can refer to multiple Positions over time. For example, a given player A that was placed in the segments 1-1, 1-1-1, and 1-1-1, in time-spans 1, 2, and 3, respectively, referenced the segment 1-1 once and the segment 1-1-1 twice. Therefore, Position 2 was referenced once and Position 3 twice. Given that there are many players, and each player can play for many days (many time-spans), the number of references each Position receives can vary. By counting these references occurrences, it is possible to identify the most common deepness of a subset. Figures 64, 65, and 66 show the number of references for each Position regarding the subsets Training, Test_1, and Test_2, respectively.

Note that the most referenced Positions for the Training, Test_1, and Test_2 subsets regarded Positions 6, 5, and 48, respectively. It means that, in the perspective of deepness, the Training and Test_1 subsets present a similar behavior, whereas the Test_2 differs from them. Another difference between the subsets is that they vary in the number of possible Positions, where the Training has 33, the Test_1 35, and the Test_2 115. These differences can be justified by the change in the game business model, from subscription (Training and Test_1) to free-to-play (Test_2), where a greater amount of new players started to play, presenting a distinct behavior.

This dataset provides for each player a churn label (which will be used in the next Chapter to assess the proposed psychological profile). Given this, it was proposed to verify if there is a Game Path that presents a distinct churn occurrence, highlighting the opportunity to mitigate the churn problem from a proactive perspective, where counter-measures can be applied to prevent churn even before the players start to lose motivation in continuing playing. This computation consisted of ranking all the Game Paths Segments based on

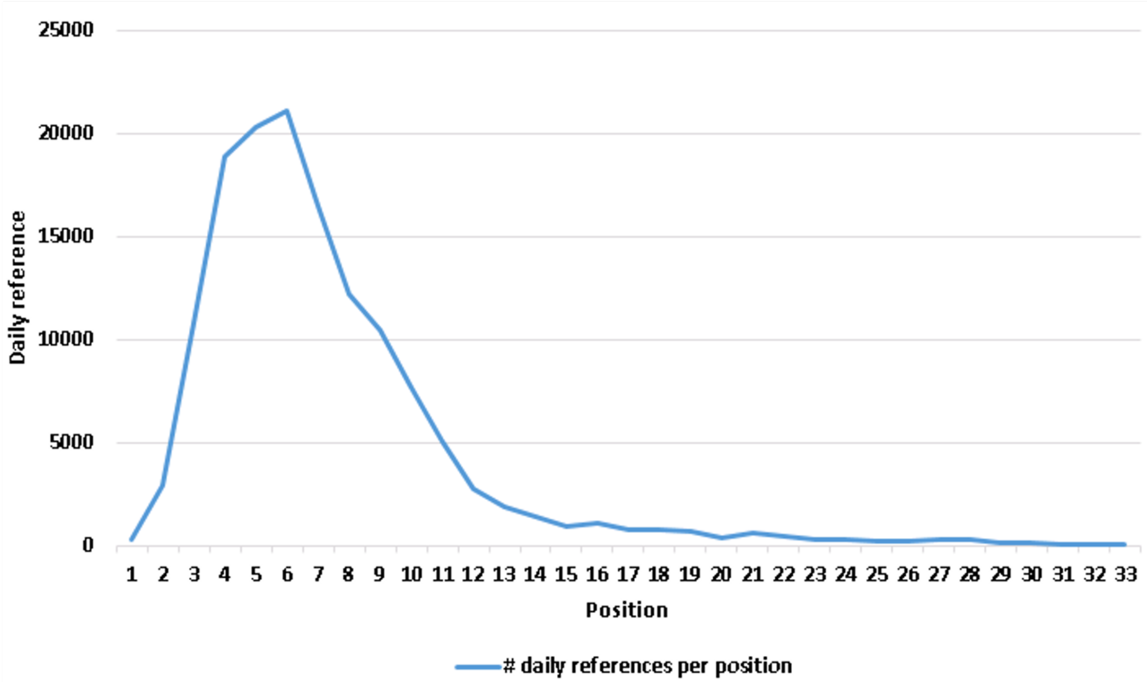


Figure 64 – The daily reference per Position of the Training subset

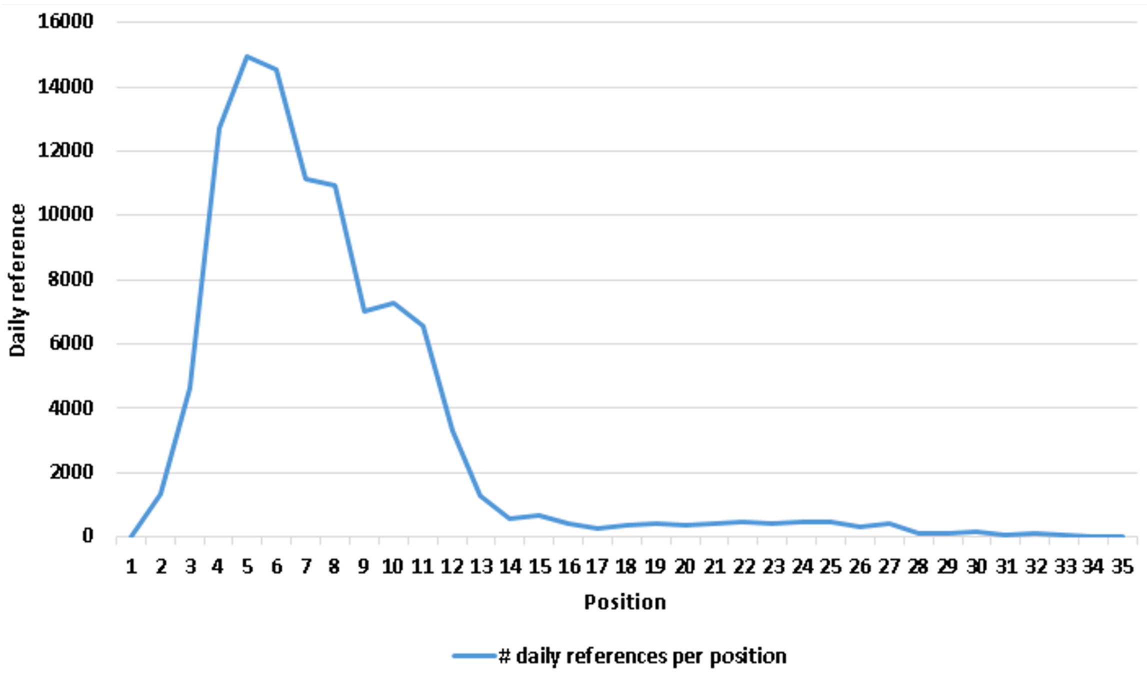


Figure 65 – The daily reference per Position of the Test_1 subset

their number of churners. Interestingly, such a Game Path was identified, regarding the Game Path 23, which, hereafter is named "Churn Path"[8]. In this Churn Path, even though splits are performed, and consequently increasing the specificities of the players' choices,

8 This Churn Path regards 114 different Positions and starts at the Game Path Segment 23 and goes until the Game Path Segment 23-2-1-1-1-1-1-1-1-1-1-1-1-1-1-2-2-1-2-1-2-1-1-1-1-1-2-2-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-2-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-3-3-2.

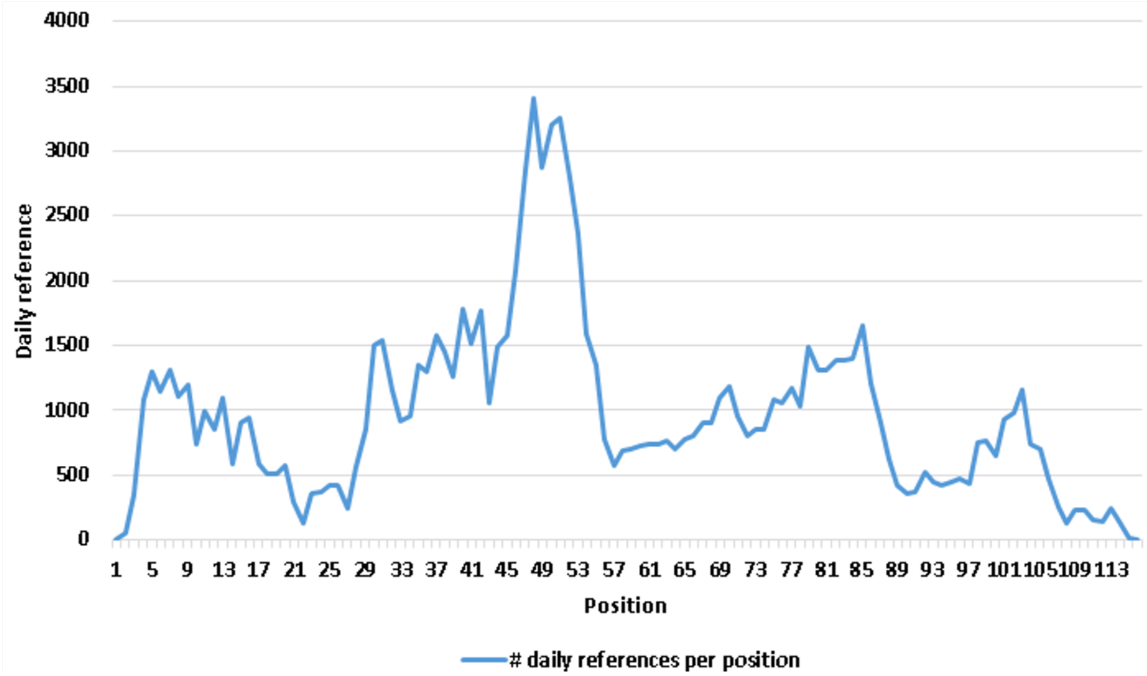


Figure 66 – The daily reference per Position of the Test_2 subset

there is a "main road" of churners, with 114 Positions, where the more players "walk" into this road, the greater is the number of churn occurrences. To compute the churn probability to each segment, Equation 6.1 was adopted, where $TotalOfChurners_i$ regards the number of players with a positive churn label in the segment $_i$, and $TotalOfNonChurners_i$ the number of players with a negative churn label in the same segment $_i$. Figure 67 presents the churn probability of this Churn Path (note that the value of 1 at its last Position regards only one player). In addition, Figure 68 shows the number of churners and non-churners present on each Position of this Churn Path. It is essential to highlight that, as soon as a player is placed in this Churn Path, and even though this player is entertained, he/she has a 30% of a chance of churning, which demands counter-measures by the game producer to try to maintain this player. The advantage is that such counter-measures are proactive, not being required to wait until a player start to feel demotivated to act, which means a change of perspective in the churn management from reactive to proactive. An interesting fact, is that the first 31 positions of the churn ranking regarded this Game Path 23, fomenting its interesting nature of portraying risk situations even before they occur (i.e., players' churn) (as this ranking has 15,861 elements, it is not presented).

$$\mathrm{ChurnProbability_i} = \frac{\mathrm{TotalOfChurners_i}}{\mathrm{TotalOfChurners_i} + \mathrm{TotalOfNonChurners_i}} \tag{6.1}$$

Moreover, even though the Game Paths allow a proactive approach, it is also possible to perform a reactive approach. It can be done by looking at the Game Path Segments that present elevated percentages of churn. Such as shown by the ranking in Figure 69, where only the segments with more than ten players are considered (encompassing 400 segments

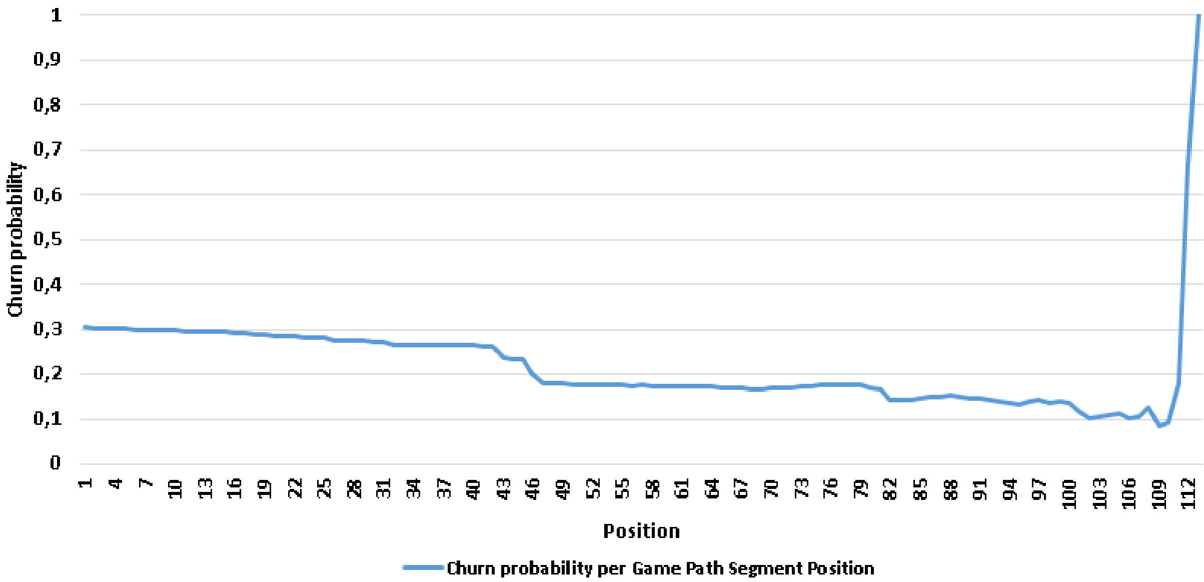


Figure 67 – The Churn Path churn probabilities

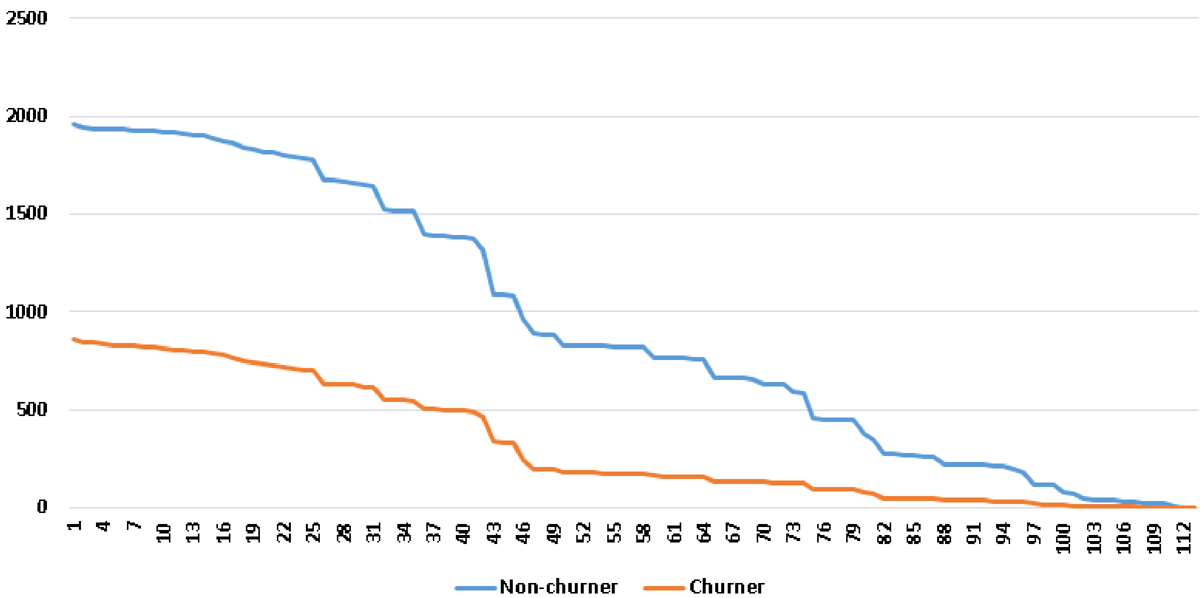


Figure 68 – The Churn Path numbers of churners and non-churners

of the total of 15,861). As we can notice, the churn probability varies from 81.25% until 13.3%. Also, by attaching to this analysis the number of players linked to each element of the ranking, as presented in Figure 70, it is possible to identify a tendency, where the higher the churn probability, the lower is the number of players (in general).

Given it, a game producer can act in two main ways to mitigate churning from a reactive perspective:

1. To focus on supporting the players with a greater chance of churning.

2. To focus on supporting the players with a lower chance of churning.

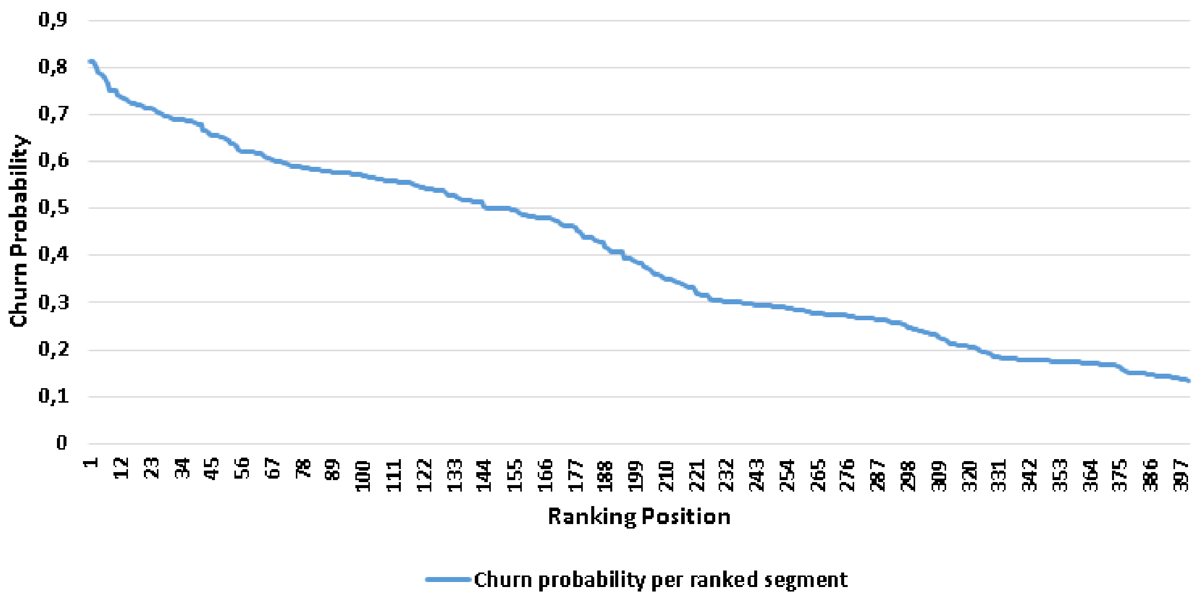


Figure 69 – Ranking of segments, with more than ten players, according to their churn probability

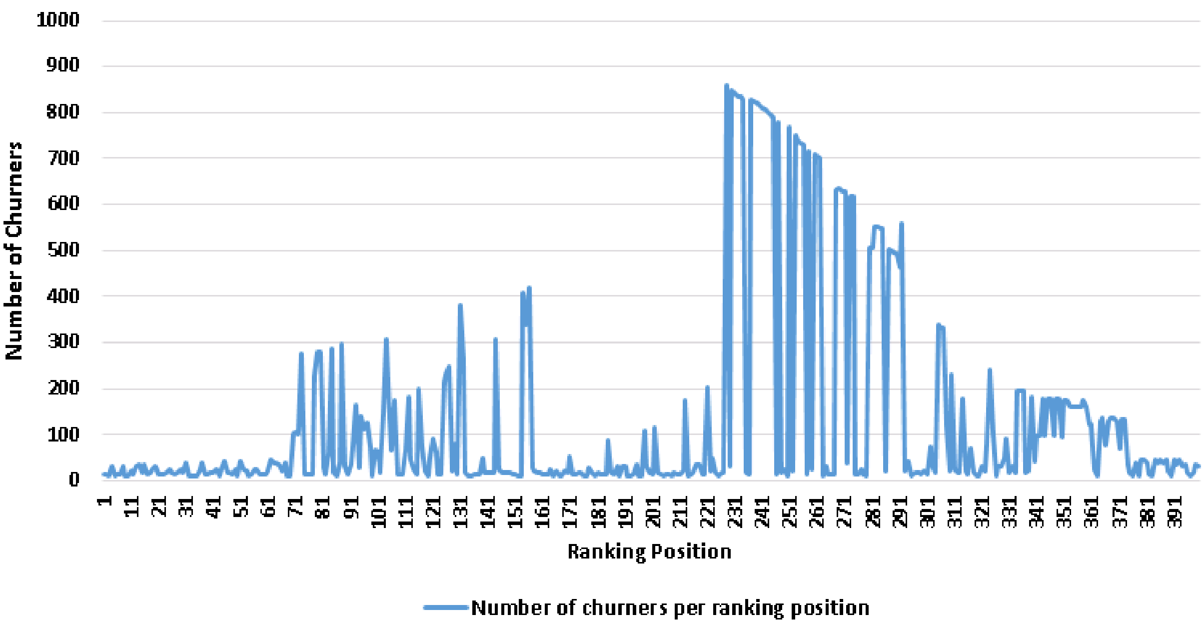


Figure 70 – The number of players per Churn Ranking placement

The first option aims to identify very unpleasant aspects of the game, and consequently, the mitigation of it can reduce the churn occurrences; however, to fewer players compared to the second option. The second option aims to cover a greater number of players; however, as the churn probability is lower, the identified unpleasant aspects are less significant than those of the first option. In sum, it is suggested that a game producer executes both options, entailing in mitigating the churn problem from both perspectives, where one looks for very unpleasant aspects while the other less unpleasant ones. Alternatively, a game company can choose not to focus on any of these options and consider only the “whales”, regardless of their churn probabilities.

A final measurement present in the Micro Spectrum regards the amount of available content. Figures 71, and 72 show, respectively, the total amount of available content and its percentage per player for all subsets. Following the findings related to the sharings of segments, the available content is also low. In the best case, the first day, an amount of 41 Game Path completions were available for players to consume the game content in the same way as other players did. Also, by dividing this amount for each player in a percentage perspective, the greatest value was 0.856%. Looking from the last day point of view, these lower values are explained by the fact that only two players are placed on the same Game Path Segment as another player, highlighting that almost all players are the heads of their segments, which means that they cannot consume contents in the same way as other players did, entailing lower available content. The definition of an ideal value of available content to define when a game upgrade should be released is still an open question, but we suggest that when the available content is continuously decreasing, it is an alert, and game producers should consider a release to keep players feed with new content.

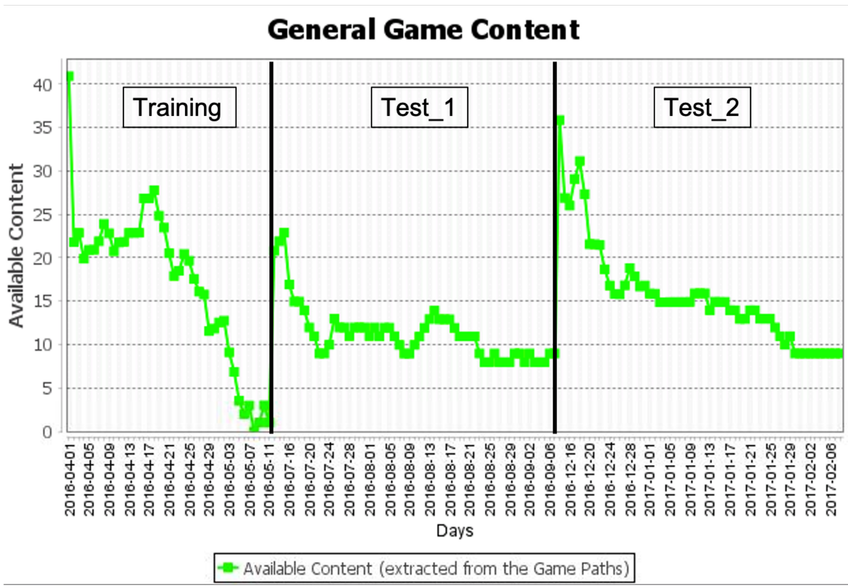


Figure 71 – Available content for all subsets

6.3.2.2 Macro Spectrum Results

The Macro Spectrum analysis is simpler than the Micro one, as only the chase pattern of human needs is considered. As initially depicted in the Sentiments Results Section 6.3.1, it was identified that players prioritize different needs, as well as do not chase some of them. Figure 73 complements the sentiment findings of player 0021D8AA (Figure 43) according to his/her Macro Distribution. As we can see, even though the Materialism does not have the highest sentiment value, it is the most chased need to this

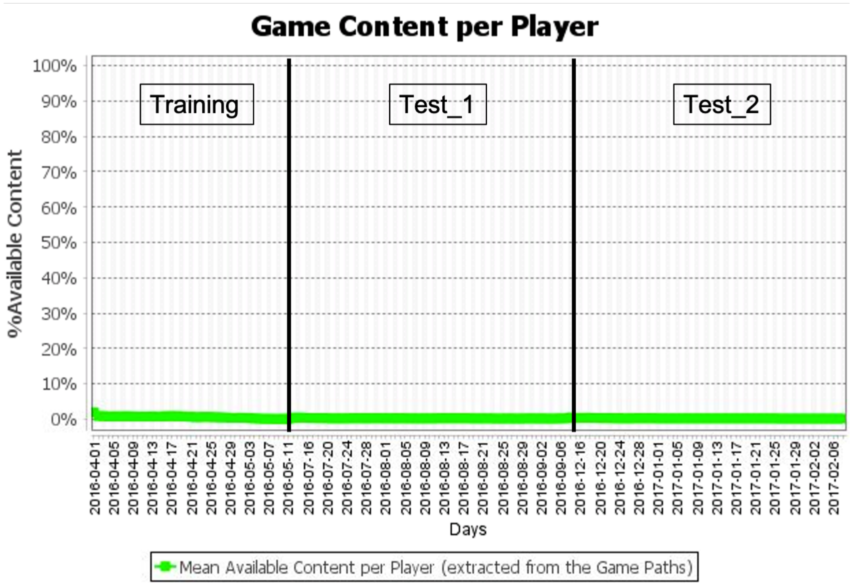


Figure 72 – Available content per player for all subsets

player. Note that this linked analysis of sentiments and the Macro Spectrum provides insightful findings, such as the linking between identifying what a player most chases and his/her success rate in doing it.

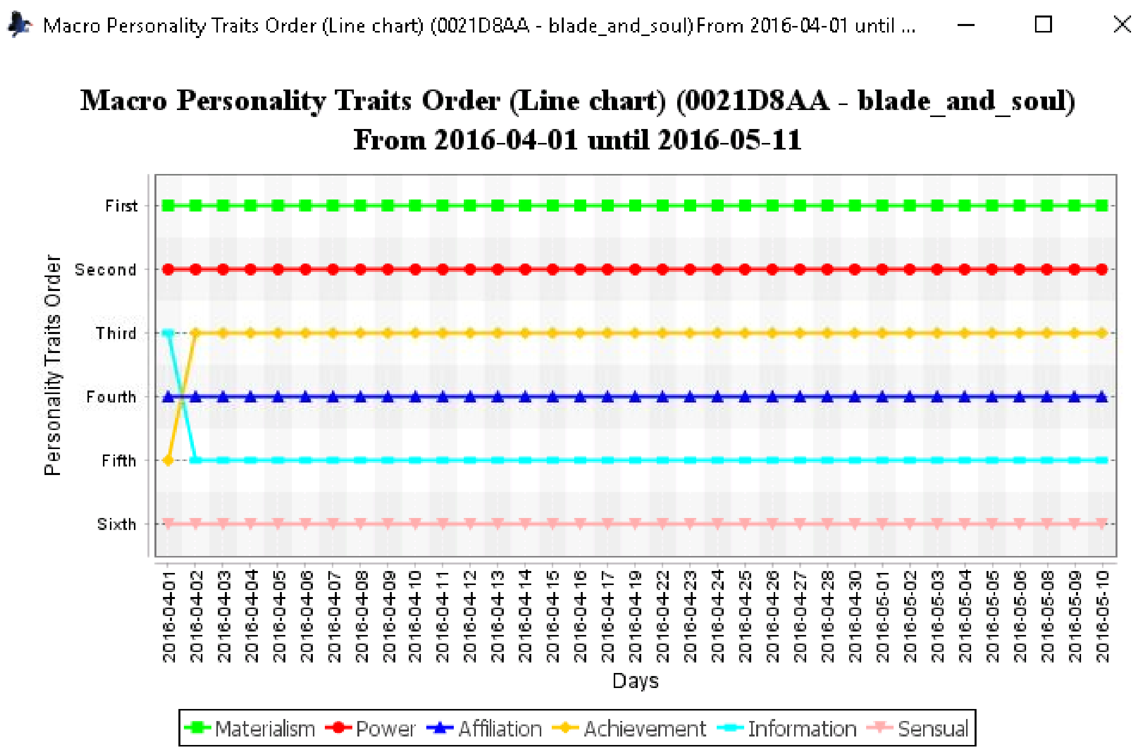


Figure 73 – Player 0021D8AA Macro Distribution

Moving to a general point of view, Figures 74, 75, 76, 77, and 78 depict the percentage of players that prioritize each need as the first, second, third, fourth, and fifth most chased, over time.

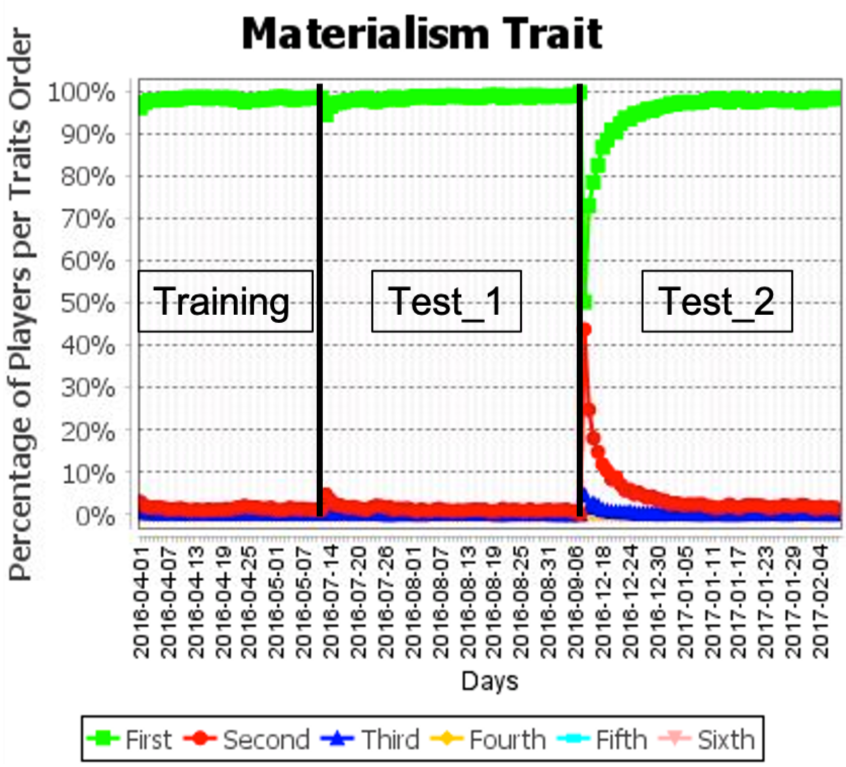


Figure 74 – Materialism chase priority for all subsets

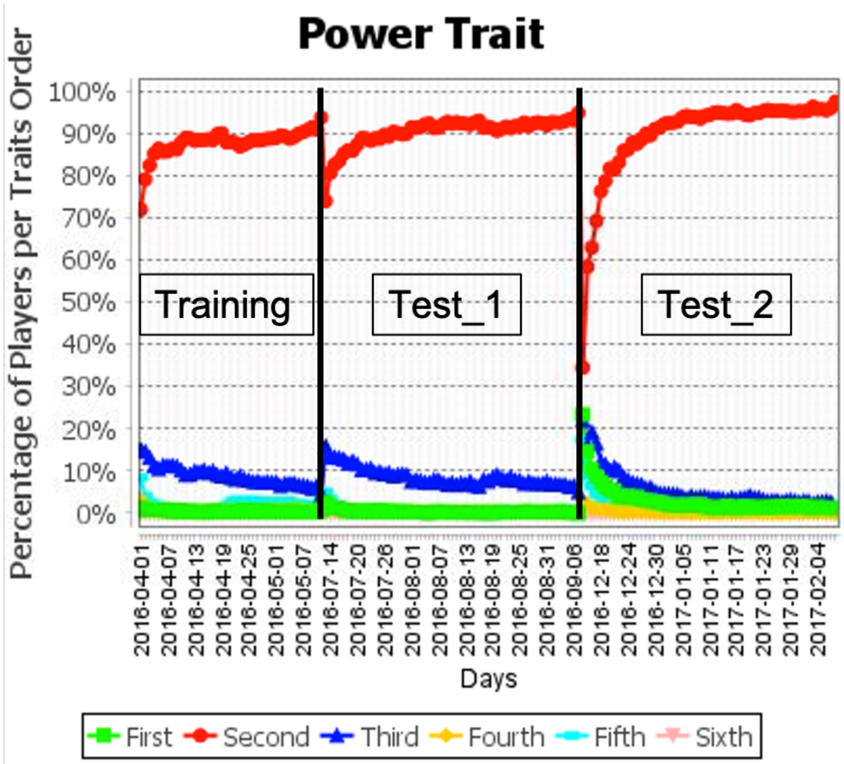


Figure 75 – Power chase priority for all subsets

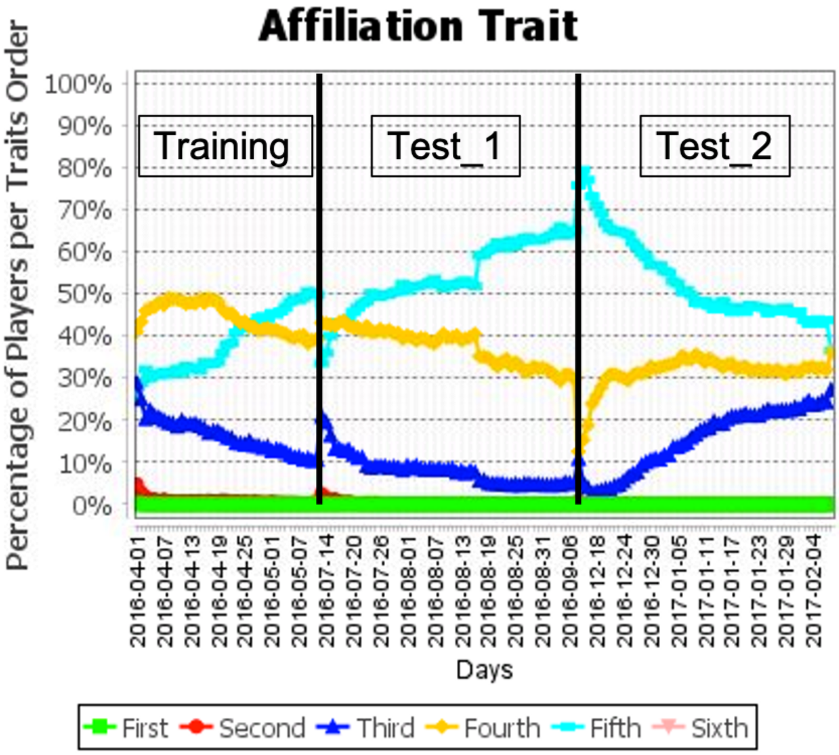


Figure 76 – Affiliation chase priority for all subsets

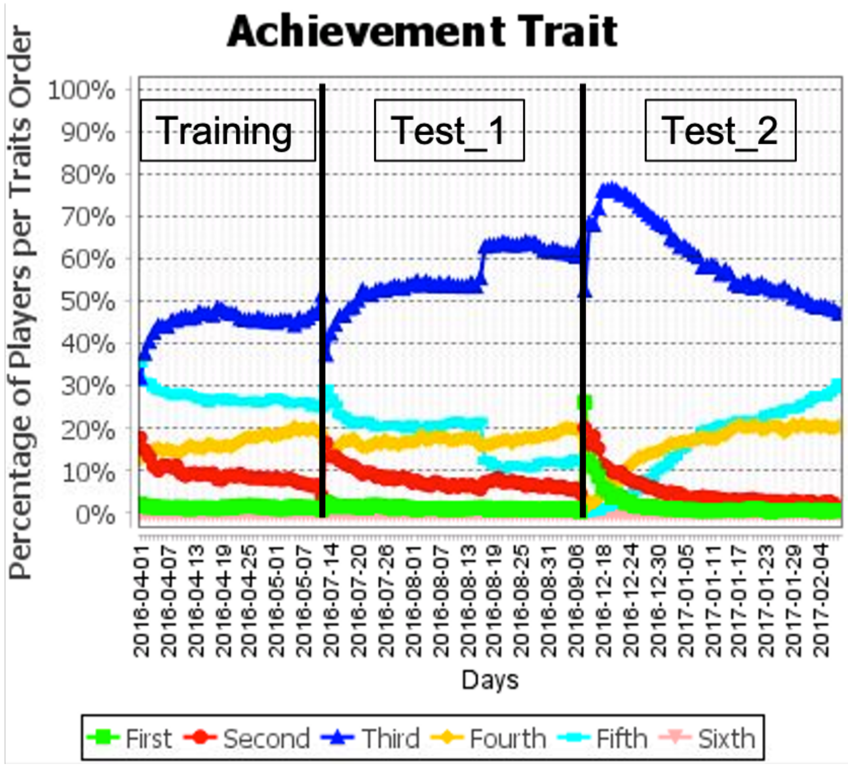


Figure 77 – Achievement chase priority for all subsets

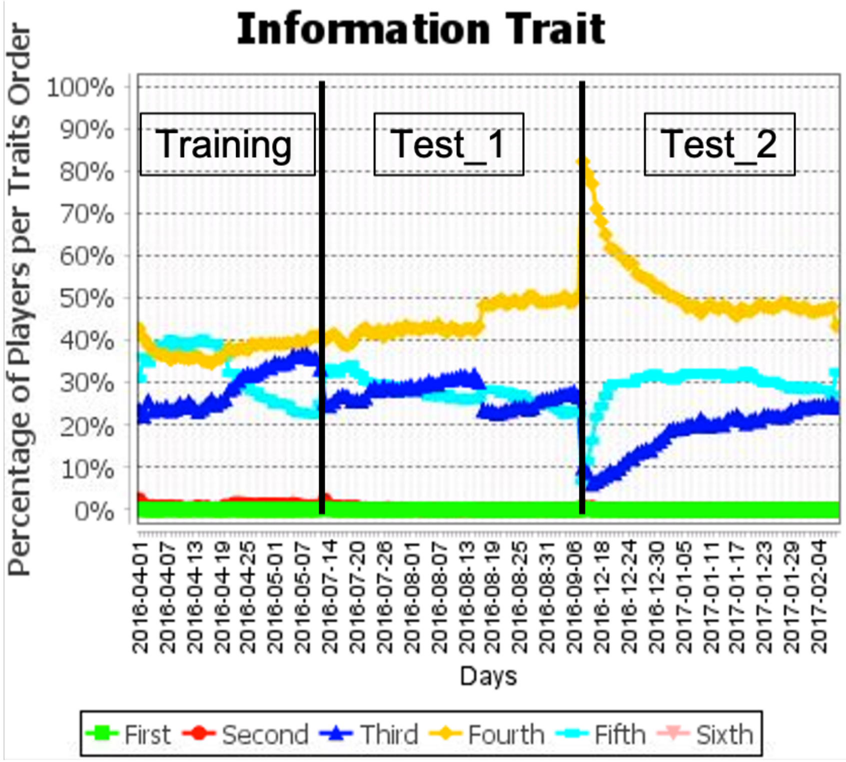


Figure 78 – Information chase priority for all subsets

As previously mentioned, few players are present on more than one subset (as depicted by Table 70); thus, each subset presents its own behavior. Given these arrangements, it is possible to suggest the following rank (Table 72) that points out the most common chased needs.

Table 72 – Ranking of the most chased needs of the Macro Spectrum for all subsets

Position	**Human Need**
1st	Materialism
2nd	Power
3rd	Achievement
4th	Information
5th	Affiliation
6th	Sensual (not offered)

Interestingly, all needs presented a similar ranking over all subsets, except for some sporadic cases in the Training subset, meaning that the Table 72 depicts the pattern of content consumption to this game. This means that an upgrade that affects the Materialism need (1st position) impacts players more than an upgrade that deals with the Affiliation need (5th position).

In addition to the depicted ranking of needs, Figure 79 shows the percentage of needs chased and not chased. Note that, except for the Sensual need that is not offered by the game content (100% of not chase), the Affiliation need is the most not chased need, with a percentage of 3%. Linking this fact with the Table 67 descriptions, where the Affiliation need is the one with the highest possible interactions (27), this 3% represents the identification of a latent and relevant social aspect, as these players are unsocial by avoiding more than 50% of the available actions that attain or impair human needs.

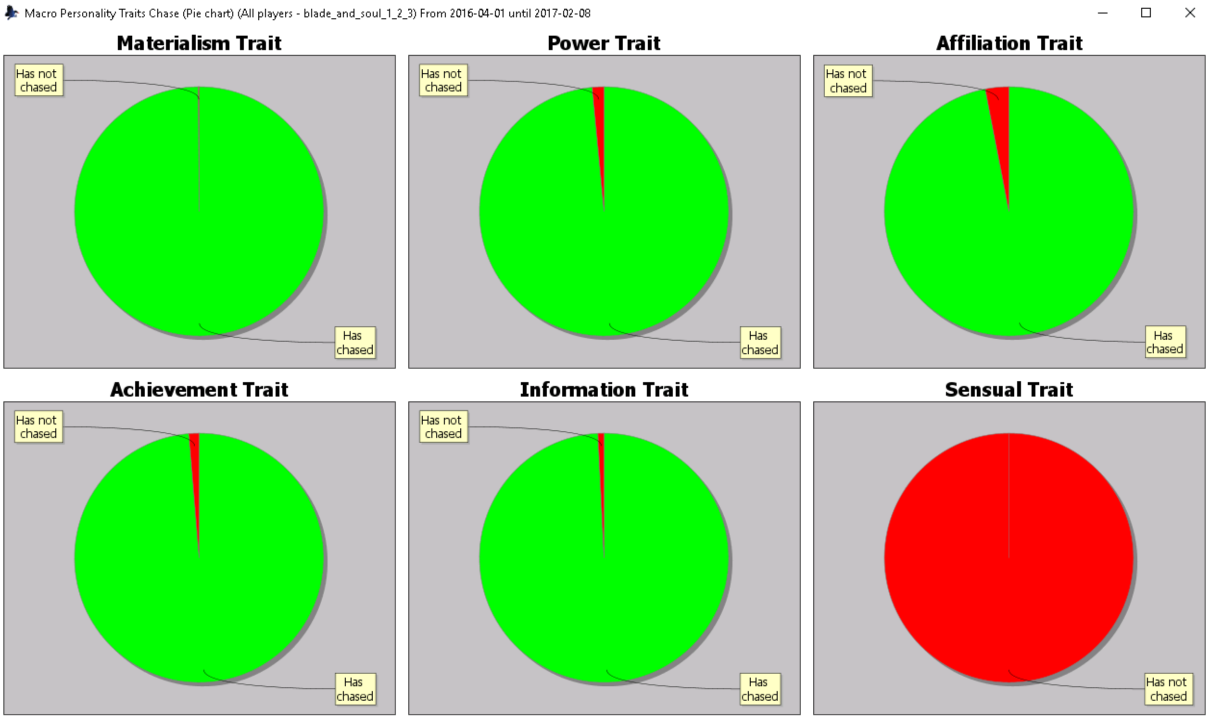


Figure 79 – Chase and not chase of human need for all subsets

In conclusion, it is essential to highlight that a player's Macro personality regards his/her choices (i.e., what are his/her priorities, what he/she likes to chase, and what he/she does not like). Also, the same player can present variances on their personality ranking according to the game content available to him/her. It means that, for the same player, some needs can change position over time, as well as not chased needs being chased; however, it is expected that the longer a player plays, the more stable his/her Macro personality is (such as depicted in Test_2 subset, where each need tends to stabilize over time).

6.3.2.3 Micro and Macro Similarities Results

The Micro and Macro Spectrums provide means of identifying a player's personality by considering all the identified choices since his/her first action in-game (i.e., the idea of a long-term aspect). Nevertheless, it is expected that the similarity between players in the Micro and Macro Spectrums differ, given that each spectrum carries its own specificities. On the one hand, the Micro Spectrum is very detailed, where a simple different sequence of actions is sufficient to distinguish two players' personalities, whereas, on the other hand, to distinguish two players' personalities in the Macro Spectrum, a change in the needs priorities is needed. Moreover, in the Micro Spectrum, once two players are different, they cannot be equal again, what does not happen to the Macro case, where players can prioritize the same needs in the same manner, regardless of whether they present a previously different ranking composition or not. In sum, it is expected that the Micro similarity is lower than the Macro one. Next, Figure 80 presents the Micro and Macro similarities for all subsets.

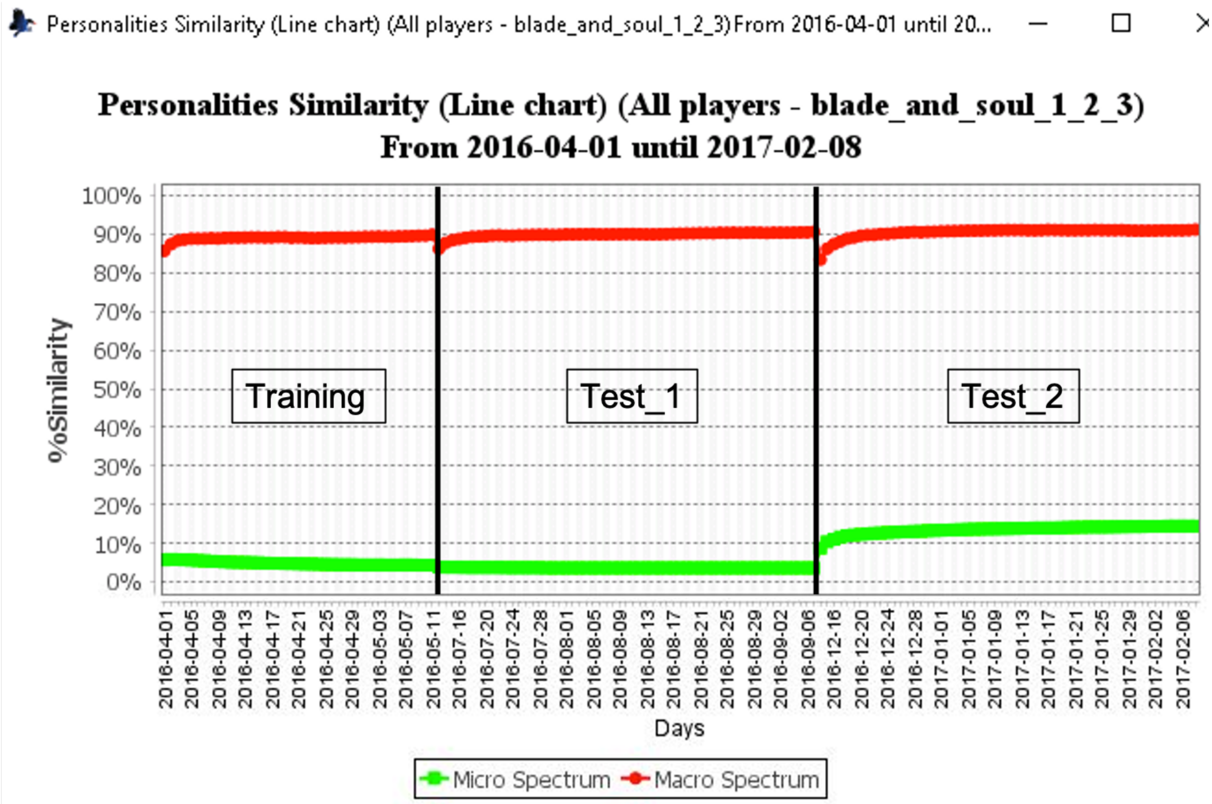


Figure 80 – Micro and Macro similarities for all subsets

On the one hand, a Macro similarity near 90% indicates that the players tend to prioritize the same set of human needs. On the other hand, a Micro similarity near 10% indicates that, even though they chase the same sets of needs, the way they chase differs in an action sequence perspective. In the Macro similarity perspective, it is possible to notice a similar behavior for all subsets. After each subset starting period, the priority chase over the human needs stabilizes near 90%, meaning that the players' priority chase achieved a stable level, where the rankings change less. In the Micro similarity perspective, the Test_1 similarity was lower compared to the one of Training. In particular, we highlight the Test_2 similarity, which was greater than the other subsets. It is explained by the fact that this subset has a greater number of new players, which were entailed by the game business model change from subscription to free-to-play. A hidden factor that foments this increased value regards a bias present in the dataset. Unfortunately, it is impossible to compute the Game Paths for all players considering all of their actions since they are not available for the old players[9]. The only cases where the Game Paths are computed without any bias are when they consider the new players, as all their actions are present in the dataset. Therefore, as the Test_2 subset has the most significant number of new players, it has the least bias, entailing the greatest Micro similarity.

As a final remark, the Micro similarity can also be seen as how similar or not the players explore the game content, highlighting aspects not shown by the Macro spectrum. For instance, how players consume a new game update.

6.4 Method Comparison

This section compares the proposed method with a parallel study with the same aim, the Game Refinement Theory. Initially, in Subsection 6.4.1, the extension in-which the Game Refinement Theory Value can be applied from a Game Analytics perspective is depicted. Next, in Subsection 6.4.2, the benefits of the proposed method are presented together with the ones of the Game Refinement Theory, highlighting possible complement points between them.

6.4.1 A Parallel Study of Players' Behavior Identification - A Game Refinement Theory Approach

The Game Refinement Theory is attached to measuring game mechanisms that affect the players' immanent accelerations in mind, which, depending on the identified values, mean pleasure in play. Given that, for every change in the game, such accelerations can also change. The experiment of this Subsection aims at gauging the changes in players' behaviors through measurements of the Game Refinement Value (GRV) in view of game upgrades present in the Blade&Soul dataset. This experiment is of interest to this thesis for two main reasons, where: (1) it highlights the extension in-which the GRV, which is a parallel approach compared to this thesis as it does not focus on psychological models, can depict players' behaviors from a Game Analytics perspective, and (2), if it is possible to propose a linkage between these two approaches to provide an enhanced understanding of players' behaviors, as approached in Subsection 6.4.2. Moreover, the GRV is analyzed together with three other raw metrics, providing more detailed descriptions about the players' reaction to new content.

After analyzing the historical upgrades in the Blade&Soul dataset, it was proposed to identify the players' motivational changes through four different usage metrics: Commitment, RI, GRV, and AMG (available motivational growth). Where the AMG metric is

[9] Players are labeled as old if, in their first action in the dataset, their levels were greater than one, indicating that there are actions not contained in the dataset.

a new one. Next, in Table 73, the summary of released upgrades during the dataset period is presented.

Table 73 – Blade&Soul upgrades summary

Date	**Upgrade**	**Subset**	**Week**	**Observations**
Mar 23rd 2016	Update 2.0 Silverfrost Mountains	-	-	Level 50 and Hongmoon Level 10; guild crafting system.
Apr 27th 2016	Update 2.1 Shattered Empire	1	5	New PvP mode.
Jun 1st 2016	Update 2.2 Vengeance Breaks	-	-	New dungeons and pets.
Jun 22nd 2016	Update 2.3 The Soul Fighter	-	-	New class (Soul Fighter).
Jul 20th 2016	Update 2.4 Shadows of the Innocents	2	8	More acts (stories).
Aug 24th 2016	Update 2.5 Desolate Tomb	2	13	New dungeons and PvP improvements.
Oct 5th 2016	Update 2.6 Ebondrake Citadel	-	-	New dungeons and item upgrade improvements.
Oct 26th 2016	Update 2.7 Beluga Lagoon	-	-	New PvP mode and new pet system.
Nov 16th 2016	Update 2.8 Midnight Skypetal Plains	-	-	New party's challenges.
Dec 7th 2016	Update 2.9 Ruins of Khanda Vihar	-	-	New quests.
Jan 18th 2017	Update 2.10 Anniversary Update	3	21	Cosmetic improvements and item drop modification.
Feb 8th 2017	Update 2.11 Wings of the Raven	-	-	New dungeons and improvements over item upgrade system.

As we can see, some upgrades are not represented in the usage data; however, we understand that they can influence the players' behavior in the observable period. This experiment will focus mainly on the four upgrades represented in the data, regarding PvP (5th-week), Story (8th-week), PvP (13th-week), and cosmetics and item drop modifications (21st-week).

After evaluating the Commitment, RI, and GRV metrics' potential, some gaps of concepts were identified. Thus, improvements to cover them were proposed. Starting with the RI metric (for more details, please see Section 3.4), it shows the variations of players' commitment over time. However, it is not clear when the variation was positive or negative (i.e., if most players improved or decreased their commitment degree). We proposed then to expand the metric's range to -1 until 1, where negative values regard the case when the majority of players decreased their commitment degree, and positive ones when the majority increased it. The updated RI Equation is depicted in Equation 6.2, in addition, Table 74 points the possible variations of commitment used by this metric.

Table 74 – Possible changes in commitment

Change Group	**Label**	**Engagement**
Low to Average	LA	Growth
Average to High	AH	Growth
Low to High	LH	Growth
Average to Low	AL	Decay
High to Average	HA	Decay
High to Low	HL	Decay

$$\mathrm{RI} = \frac{(LA + AH + LH) - (AL + HA + HL)}{max_{positive}(RI) \veebar max_{negative}(RI)} \tag{6.2}$$

Where LA, AH, LH, AL, HA, and HL are the same as depicted in Table 74. In addition, when the sum $(LA + AH + LH)$ is bigger than the sum $(AL + HA + HL)$, the $max_{positive}(RI)$ is used, being the $max_{negative}(RI)$ used otherwise [10]. In the previous version of this metric (KUMMER; NIEVOLA; PARAISO, 2017b), only the $max_{positive}(RI)$ was used.

Another gap presented in the Commitment and RI metrics regards the potential growth of commitment per time-span. As it is not explicitly presented in these metrics' concepts, we proposed to use a new one, that we called "Available Motivational Growth" (AMG). This new metric consists of a sum of the number of players not in the high commitment degree, such as depicted in Equation 6.3.

$$\mathrm{AMG} = \frac{P_{low} + P_{average}}{max(AMG)} \tag{6.3}$$

Where P_{low} and $P_{average}$ are the number of players with low and average commitment respectively, and $max(AMG)$ is the biggest AMG of the series [11]. The metric's range is

[10] Note that the values of $max_{positive}(RI)$ and $max_{negative}(RI)$ are used to normalize the Equation final result. Such normalization is computed over a series of previously computed values regarding the Equation part $(LA + AH + LH) - (AL + HA + HL)$.

[11] It follows the same normalization logic applied to the RI.

from 0 until 1, where 0 means no potential growth (i.e., all players already have a high commitment) and 1 the opposite (i.e., the best potential of the series). It is interesting to highlight that this range of values allows a combined analysis with the RI metric, as it highlights the potential RI growth per time-span independent of the RI identified value. It means that even though the RI shows a value of 1 (the best growth), the AMG value could highlight a hidden potential (i.e., the best performance could be even better). Moving to the metrics computation, Figure 81 shows the Commitment distribution associated with the game upgrades.

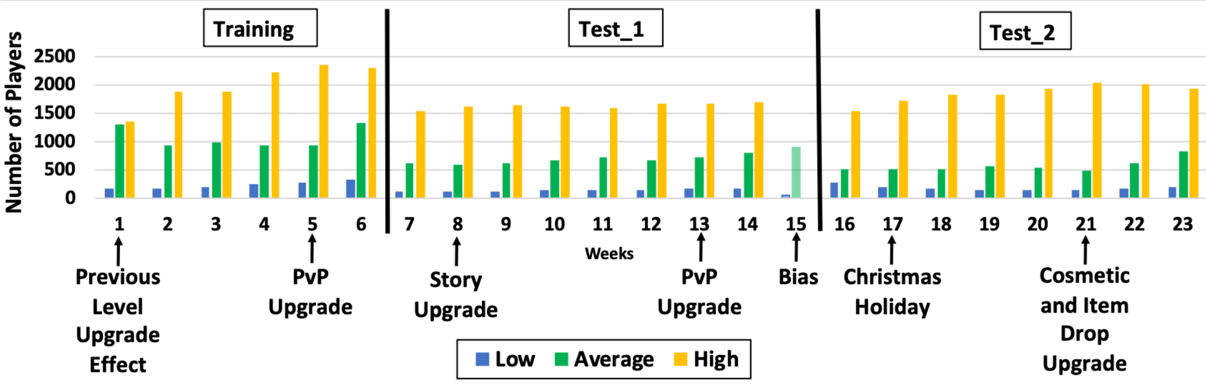


Figure 81 – Commitment assignment for Blade&Soul with upgrade labels

As it is possible to notice, the first week had a similar number of average and high committed players. The reason for that was a previous upgrade that improved the players' max level. In addition, this week can be seen as a "rush" to achieve the new highest level, entailing in a greater number of level up events per player compared to the other weeks. Moving to the Christmas holiday, it presented a growth of the high committed players. We understand this increase as the players' enjoyment due to the opportunity to play longer than usual.

A t test was applied considering $p < 0.05$ over the percentage of players on each commitment degree per subset to identify whether their distribution changed from one subset to another. The statistically significant differences identified regarded the low committed players between subsets Training and Test_1, and the average and high ones between subsets Test_1 and Test_2, and between subsets Training and Test_2. In conclusion, it was possible to identify the influences of upgrades and the change of the business model in the distribution of the players' commitment degrees, showing that different approaches can affect the measurement of players' engagement. Next, the RI and the AMG values are presented in Figure 82.

In the RI perspective, it is possible to notice that the 4th-week presents an improvement of players' commitment, which can be seen as a positive expectancy regarding the upcoming upgrade. However, it was not kept in the following two weeks. We found two possible reasons for that, (1) after the players consuming the new content, they lost

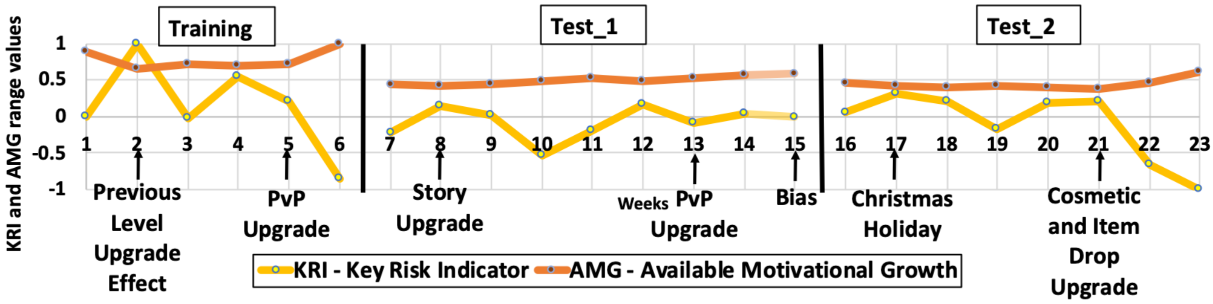


Figure 82 – RI and AMG assignments

motivation, reducing then their commitment degree to the game, or (2), their expectancy was not attained. In the next upgrade (8th-week), the players' commitment grown compared to the previous week, but again, it dropped in the following weeks due to game content consumption or expectancies not attained. The 13th-week and 21st-week upgrades can be considered unsuccessful because the players' commitment dropped; also, the 21st-week presents the worst decay of players' engagement, highlighting possible disapproval about the new content.

Even though the level upgrade week is not present in the dataset, its effects could be measured two weeks after its release as the best improvement of players' commitment (2nd-week), as many players improved their level together with an increased usage time. We can also conclude that this kind of upgrade has the best potential to improve the general players' commitment to an MMORPG.

Moving to the AMG perspective, none of the upgrades was engaging enough to have an AMG value lower than 0.4. The highest value was identified in the week after the first upgrade (6th-week), highlighting the drop of players' engagement in this time-span. Additionally, the best RI value happened together with an AMG value of 0.66. It means that even though it was the best growth of players' engagement, there was a hidden potential to get better results. As a final remark, the AMG mean value was 0.55, which means that there was an opportunity to captivate players to higher degrees of commitment over the whole dataset period, highlighting the opportunity to promote game upgrades and the difficulty to motivate different kinds of players at the same time with them.

Turning to the GRV perspective, we could apply it to two game mechanisms: the PvP battles and the Reforging systems (i.e., an item upgrade system). For the PvP case, we used the number of fought battles as D and the number of won battles as B, while for the Reforging case, the number of attempts was assumed as D while the number of successes as B (for more details about the computation of the Game Refinement Value, please see Section 3.5). Next, Figure 83 shows their results attached to the players' commitment degrees.

Firstly observing the PvP perspective, it is possible to see that the identified GRVs

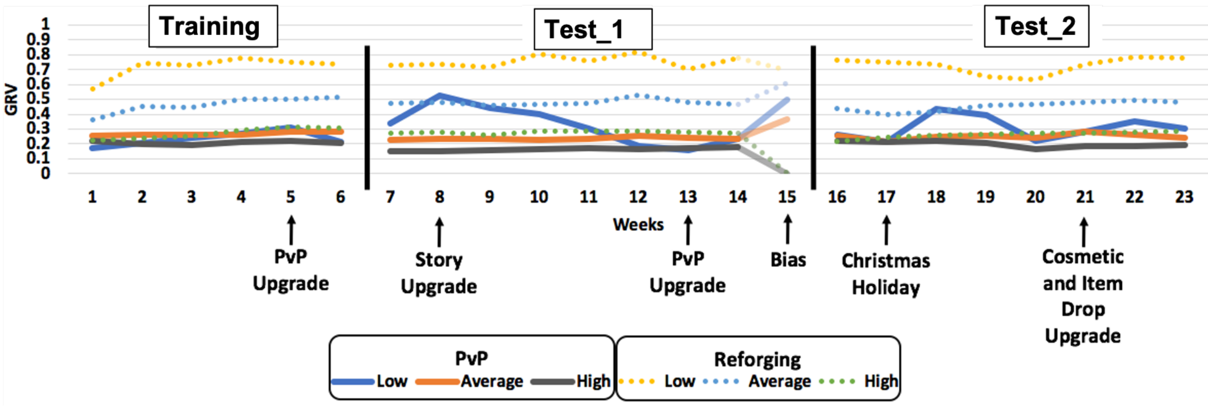


Figure 83 – PvP and Reforging GRVs assignments

are not in the called "sophisticated zone" (i.e., the range between 0.07 and 0.08; the also called "noble uncertainty" (YICONG et al., 2019)). After assimilating the zone's concepts to the current data, the PvP battles can be seen as unfair and based on chance due to their high GRVs. However, we understand that it is not entirely accurate, as the PvP battles in the considered game have a different aspect attached to the players' skill that we named "evolutionary strength". During a battle, the players' skill can be seen as the players' effectiveness and efficiency in using the options available to them (i.e., "playing well"), while the evolutionary strength is the players' acquired level until a given moment. In addition, the evolutionary strength is accumulative; thus, once it is acquired, it is not possible to lose, what does not happen to the players' skill, which may improve with training or reduce due to a lack of practice. For example, the combination of such characteristics can generate situations where a very skilled player may lose a battle against a less skilled one if the other player's level is higher. We understand that this kind of situation influences the concepts around the sophisticated zone. For example, the fact of a player having more levels than the other does not mean that the battle is based on chance; therefore, we considered the PvP of the Blade&Soul game as possibly unfair and based on a combination of players' skill and evolutionary strength. As a final remark, statistically significant differences were found in a t test with $p < 0.05$ considering the set of GRVs per commitment degree, the only exception was between the low and the average groups.

Moving to the Reforging perspective, it was not possible to distinguish the effects of the different kinds of upgrades to any commitment degree, a possible explanation being the lack of improvements in the Reforging system during the dataset period. However, as occurred in the PvP battles, all the GRVs identified are not in the sophisticated zone. A reason for that regards the general success rate of approximately 94% that reduced the sense of outcome uncertainty. Moreover, all commitment degrees presented statistically significant differences with $p < 0.05$. Thus, according to the identified values, we consider the Reforging system of Blade&Soul as entertaining and based fully on chance.

As general conclusions, the RI indicates that PvP upgrades tend to improve the players' commitment before the upgrade, presenting then a drop of engagement after the new content consumption, while story upgrades can keep players motivated for longer compared to the PvP ones, presenting picks of motivation during the upgrade week. Moreover, the upgrade regarding cosmetic and item drop modifications presents an initial growth before the upgrade and a substantial decay after it, similar to the PvP one. According to these behaviors, we could point three reasons where: (1) new PvP battles are opportunities to players became in evidence by winning the most battles as possible, thus an initial preparation is a good idea to have a good performance when the new competition starts, moreover, the competitions seem to excite or demotivate players fast (i.e., one week); (2) there is no preparation to accomplish story upgrades, therefore players become more committed only when the new content is available by accomplishing the new challenges; and (3) upgrades regarding cosmetic and item drop modifications generates an initial expectancy or preparation, but after a player acquiring his/her new costumes or desired items, his/her engagement drops as the objectives were accomplished.

In the Game Refinement Theory perspective, the higher a GRV, the more entertaining a game is, while the lower, the more competitive. Thus, it is possible to state that high committed players present a more competitive behavior than the lower committed ones. On the one hand, considering the high committed ones, even though there were different kinds of upgrades, their GRV keeps similar. We rationale two possible reasons for that, where (1) the degree of challenge offered to them keeps the same over time, or (2) the players could adapt and maintain their performance. While on the other hand, players in lower commitment degrees have changes in their GRVs. In conclusion, low and average committed players present more changes in their sense of entertainment considering PvP battles and the Reforging systems over time than the high committed ones; or in other words, low and average committed players are more sensitive to changes.

In summation of all findings regarding the four considered metrics, we could identify five main characteristics behind the players' motivational changes, being them: (1) different game upgrades entail in expectations to consume the new content; moreover, some kinds of them allow a preparation while others do not; (2) depending on the considered upgrade, players can increase their commitment to the game for longer; (3) upgrades that present a preparation aspect tend to keep players engaged for less time than the ones that do not allow it; (4) the more committed a player is, the more resilient he/she is to changes on game mechanisms; and (5) the higher the commitment of a player, the more competitive his/her behavior is.

6.4.2 Comparison Between the Proposed Method and the Game Refinement Theory

The proposed method and the Game Refinement Theory share the same aim, to better understand players' behaviors; however, each approach focuses on a different nature. On the one hand, the Game Refinement Theory is based on physics concepts applied to the human brain, whereas, on the other hand, the proposed method is based on textual descriptions of players' behaviors. Moreover, both approaches require data as input. Therefore, the challenge is to find a way to join both perspectives to provide an enhanced understanding of players.

From the proposed method perspective, its benefits encompass the translations of players' actions in-game to psychological metrics that cover short, mid, and long-term aspects. Moreover, the resultant psychological profile comprises two parts, the quantitative metrics[12] and their respective GCs of the unified human-being model. It means that given changes to these metrics values, the interpretation of such changes can be supported by GCs' textual descriptions. In sum, the proposed method provides means for game producers to interpret their players' behaviors through a lens that highlights what is happening to their players, encompassing what players desire and how successfully or not they are in achieving their wishes.

Moving to the Game Refinement Theory perspective, it regards the modeling of qualitative aspects of games based on physics concepts applied to the human brain, where depending on the identified GRV, assessments can be taken (e.g., regarding the "sophisticated zone"). The GRV can also be analyzed from two perspectives: a general view of a game considering all its players or an individualized view of a player. In addition, these views can be segregated by game mechanism (such as happened to the PvP and Reforging system in Subsection 6.4.1). In sum, the benefits of the GRV regards the identification of qualitative metrics of games, which game producers can take into account to interpret its players, as well as propose adjustments to its game's mechanisms to reach a more desirable balance between entertainment, challenge, chance, and fairness.

Bearing in mind the requirements and the benefits of each approach, it was concluded that they are not exclusive but complementary. On the one hand, the proposed method presents quantitative metrics regarding six types of human needs, whereas, on the other hand, the Game Refinement Theory provides qualitative measurements. Given this, an ideal joining between them would entail quantitative metrics that also carry qualitative aspects. Considering that the GRV can be computed based on the number of attempts and successful attempts, it is possible to compute it for each human need since the notions of

[12] As the proposed method does not differ attainments and impairments by any qualitative aspect (e.g., when an attainment regards a unique situation, like defeating the most challenging enemy), it is assumed that all generated psychological metrics are quantitative.

attempts (i.e., the sum of impairments and attainments) and successful attempts (i.e., only the number attainments) exist in such a context. With this, both original approaches are kept the same; however, their joining opens space for the computation of new and enhanced metrics, where each human need would have a new metric carrying the qualitative notion of acceleration in mind, the idea of "psychological acceleration in mind". Interestingly, such a joining has the potential to identify a kind of "human need based sophisticated zone", enhancing our comprehension of players' behavior. Note that such a joining does not require additional information (i.e., the input is the same as the proposed method one).

7 Method Experimental Assessment

This Chapter aims at assessing the generated psychological profile of players in a churn prediction problem, where if the resultant features support a better prediction of churn, the proposed profile is assessed as accurate, at least, to the churn context. This Chapter is organized as follows: in Section 7.1 the comparison baseline is presented, followed by Section 7.2 where the experimental protocol is explained, after, in Section 7.3, the experimental results are shown, highlighting if the proposed psychological profile provides a better representation of players' behaviors, or not, based on its performance against the baseline, and finally, in Section 7.4, given the results, the term "Psychological Data Enhancement" is coined to reference the general idea of the proposed method.

7.1 Comparison Baseline

The baseline regards the best churn prediction result applied to the CIG 2017 - Game Data Mining competition dataset (LEE et al., 2018) (i.e., the Blade&Soul dataset adopted in this thesis), which is the work of (KUMMER; NIEVOLA; PARAISO, 2018a). On their approach, a novel feature engineering technique was applied to generate enhanced features that carry the notions of temporal tendency.

Moreover, the Data Mining competition splitted the churn prediction into two challenges, a boolean prediction (i.e., churner or non-churner), and a survival time prediction (i.e., the number of days each player will continue playing)[1]. For this thesis, we are focused on the boolean prediction. The competition rules were the following ones:

1. Three subsets are provided, Training, Test_1, and Test_2. However, only the Training subset has the churn labels.
2. There are no restrictions regarding any algorithm or feature engineering approach.
3. A classifier should be induced based only on the Training subset.
4. The induced classifier has to be applied to the Test_1 and Test_2 subsets [2].
5. The approach with the best mean F-1 score is the winner[3].

[1] The baseline obtained the best results for both challenges.
[2] Note that the Test_1 and Test_2 subsets must receive the same feature engineering process applied to the Training subset.
[3] The mean F-1 score is computed based on the F-1 scores of the Test_1 and Test_2 subsets.

The baseline work approached four different classifiers, and their obtained results are presented in Table 75 (all results were obtained with the algorithms default configurations and a ten fold cross-validation). Given the "Final Mean Value" of 0.706375, if an approach that considers the same set of classifiers (following the same configurations and evaluation approach) obtained a better value (i.e., a greater F1-score), it will be assumed that the baseline was overcome. In conclusion, the psychological features of the proposed method will be provided as input (after a preprocessing) to these classifiers, and, if their prediction results were better than the baseline one, the psychological profile is assessed as accurate, at least, to the churn prediction problem.

Table 75 – Churn prediction Baseline

	Training		**Test F1-Score**		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.90%	0.757	0.712	0.738	0.725
RepTree	75.68%	0.751	0.721	0.73	0.7255
MLP	71.95%	0.715	0.648	0.655	0.6515
SVM	76.93%	0.753	0.721	0.726	0.7235
			Final Mean Value		0.706375

7.2 Experimental Protocol

This Section describes the considered psychological features (Subsection 7.2.1), their preprocessing that uses LSTMs to encompass temporal aspects (Subsection 7.2.2), the LSTMs hyperparameters and architectures (Subsection 7.2.3), and the Experimental Procedure (Subsection 7.2.4).

7.2.1 Considered Psychological Features

The psychological features adopted in this assessment are grouped by psychological aspects, such as in the following arrangements:

- Sentiments
    - The six daily sentiments
    - The six historical sentiments
    - The six Hope emotions
    - The six Fear emotions
    - The six Distress emotions
    - The six Joy emotions

- Game Path (Micro personality)

  - A set of 12 one-hot encodings[4] (where each one has 33 positions) regarding the Training subset maximum Position value, which is 33. The LSTMs architectures require this data transformation. For instance, assuming a maximum Position value of 3 and a player placed at the Game Path Segment 1-1-3, his/her segment representation in a one-hot encoding format is: (1,0,0), (1,0,0), (0,0,1)[5].

- Macro Personality

  - The percentages for each human need in the first position (6 features)

  - The percentages for each human need in the second position (6 features)

  - The percentages for each human need in the third position (6 features)

  - The percentages for each human need in the fourth position (6 features)

  - The percentages for each human need in the fifth position (6 features)

  - The percentages for each human need in the sixth position (6 features)

- Similarity

  - The Macro Similarity

  - The Micro Similarity

  - The general available content

  - The mean available content per player

- All features

  - It comprehends all the 472 features of the previous groups[6].

It is essential to highlight that each one of the aforementioned features was computed for each player and for each time-span, allowing in that way a time-series analysis to identify temporal aspects (such as performed in the preprocessing depicted in the next Subsection 7.2.2). In particular, the emotions of Hope, Fear, Joy, and Distress were added to the Sentiment group because they carry additional aspects compared to the other emotions, regarding the players' expectancy to each time-span. Moreover, the Similarity group is the only one to consider (together with the "All features" group), besides individualized aspects of players, game-related metrics, such as the general available content.

[4] This value of 12 was identified through experiments and represented the best Game Path deepness for the Training subset to identify churners, such as shown further in Table 78.

[5] Note that there are nine features to represent this player's Game Path Segment placement and that the higher the maximum Position value, the more features are needed.

[6] Note that 396 features regard only the one-hot encodings of the Game Paths (twelve Positions where each one has 33 values).

7.2.2 Preprocessing

For each group of features, a preprocessing is performed. The preprocessing consists of providing a group's sets of features to an RNN (LSTM, in this case) together with the actual churn label of each player, resulting in a list of churn labels for each group, such as shown by Table 76.

Table 76 – Resultant psychological feature engineering together with the actual value

Player ID	**Sentiment churn label**	**Game Path churn label**	**Macro Personality churn label**	**Similarity churn label**	**All features churn label**	**Actual churn label**
1	Yes	No	Yes	No	Yes	No
2	No	Yes	No	Yes	Yes	Yes

It means that, for instance, all the sentiments regarding each human need will be translated to a unique feature, named "sentiment churn label", which contains only "Yes" or "No" values. This process is performed to each group of psychological features, having as a final result the following set of features: sentiment churn label, Game Path churn label, Macro Personality churn label, Similarity churn label, and All features churn label. It is essential to highlight that, as these labels were generated on an LSTM network, they encompass temporal aspects.

By performing different combinations of the aforementioned churn labels, it is possible to assess if the psychological features are complementary to each other or not. Moreover, it is possible to assess if these psychological features (five) are complementary or not to the raw features (119) of the baseline. It is essential to highlight that the proposed psychological profile of players can be assessed as accurate from two perspectives, where: (1) the addition of the psychological features to the ones of the baseline presented a better performance, or (2) the use of only the psychological features presented a better performance.

To perform this preprocessing, first, it is needed to identify an optimal LSTM configuration regarding hyperparameters and architectures. Therefore, the next subsection provides definitions and justifications regarding each adopted configuration.

7.2.3 LSTMs Hyperparameters and Architectures

Before presenting the hyperparameters, it is essential to highlight that the Blade&Soul dataset is not balanced (i.e., the number of churners and non-churners differ), which guides some definitions. In the Training subset, the percentage of non-churners is 70%, whereas the one of churners is 30%. Note that to keep the comparison to the baseline fair, all the competition rules must be followed; thus, no information of the Test subsets is considered to obtain any advantage at any moment; therefore, their classes balancing are not approached.

All the LSTM experiments were performed using the Keras framework[7]. Given it, the hyperparameter names follow the same names adopted in the Keras. Next, each adopted hyperparameter is presented together with its value and justification. More details about each hyperparameter concept can be found in Subsection 3.6.3.3.1.

- validation_split = 0.2

  It was adopted according to the Pareto's principle (more details in Subsection 3.6.3.1).

- learning_rate = 0.001

  This rate was adopted to find a local optimal before overfitting the model. Experiments with 3.000 epochs showed that the validation loss tends to increase over time, demanding adjustments to mitigate the identified tendency to overfitting.

- momentum = 0

  The momentum was set to 0 to mitigate overfitting.

- batch_size = 10

  It was not set to 1 (online training) to avoid exclusive weights updates of very distinct behaviors.

- class_weight = 0: 1., 1: 2.33

  As the Training subset is unbalanced (70% of non-churners and 30% of churners), it was needed to adjust the class weights to these values to give the same importance for each class. Regarding the validation set, the last 20% of the train set is reserved as the validation set, presenting 70.7% of non-churners and 29.3% of churners (similarly to the unbalancement of 70% and 30% of the whole Training subset).

- loss_function = "binary_crossentropy"

  The binary cross entropy fits the boolean churn problem as it considers the error for each of the two classes (churn and non-churn).

- epochs = 3000

  After the executions of experiments, it was identified that from 3.000 epochs, the loss is near 0, indicating that the model is "memorizing" the Training set (overfitting). The following early stopping procedure helps to mitigate this problem.

[7] For more information, please see: https://keras.io/

- Early stopping regarding the val_loss (patience=200, min_delta=0.01, mode="min", restore_best_weights=True, verbose=1)

  This early stop configuration focuses on stopping the training when the validation loss has no improvements, of at least 0.01, after 200 epochs, being the weights restored according to the best epoch. A patience of 200 epochs was chosen to allow the identification of a possible new local optimal.

Moving to the architectures, the following four were proposed:

- **Architecture 1**

  First layer: LSTM with 41 units (regarding the number of days in the Training subset, 41).

  Second layer: Dense with one neuron and sigmoid output.

- **Architecture 2**

  First layer: LSTM with 41 units.

  Second layer: Dense with the number of neurons equal to the number of features in the input vector.

  Third layer: Dense with one neuron and sigmoid output.

- **Architecture 3**

  First layer: LSTM with 41 units.

  Second layer: Dropout of 20% (dropout tends to reduce overfitting by forcing the network to learn new perspectives by randomly disregarding some features between two layers. In our case, this random occurrence has a chance of 20% per connection between layers).

  Third layer: Dense with the number of neurons equal to the number of features in the input vector.

  Fourth layer: Dropout of 20%

  Fifth layer: Dense with one neuron and sigmoid output.

- **Architecture 4** (double neurons)

  First layer: LSTM with 82 (41 * 2) units.

  Second layer: Dropout of 20%

  Third layer: Dense with the number of neurons equal to twice the number of features in the input vector.

  Fourth layer: Dropout of 20%

  Fifth layer: Dense with one neuron and sigmoid output.

The idea behind these four architectures is to explore, for each group of features, how their pieces of information provide means for the RNN to divide the hyperplane between churners and non-churners, where some architectures can better fit to specific groups of features.

7.2.4 Experimental Procedure

For each group of features, all architectures are approached, where the one with the best Training F1-score is the chosen to generate the churn label for the respective group. When different architectures present the same F1-score, the one with more performed epochs will be chosen. The F1-scores of Test_1 and Test_2 will also be computed for additional analysis but not considered when choosing the best Training configuration.

In particular, we highlight the fact that the Game Path group demands additional computations to identify the ideal number of Positions to be considered (i.e., the ideal deepness). Given this, the Game Path perspective presents 33 computations for each of the four proposed architectures, where the deepness varies from 1 to 33. Also, this ideal number is adopted by the “All features” group.

After preprocessing all the groups of features and generating the churn labels for all subsets, a set of experiments explore different combinations between them and between them and the baseline’s raw features. Each experiment provides as final results the mean F1-score (of Test_1 and Test_2) for each one of the four classifiers adopted by the baseline approach, namely C4.5, RepTREE, MLP, and SVM, where all of them must be executed with the same default configuration as the baseline, considering also a ten fold cross-validation. To assess if a given combination of features overcomes the baseline, a mean value of all these mean F1-scores is computed, and if an experiment presents a value higher than 0.706375, it overcame the baseline.

7.3 Experimental Results

The preprocessing results for each group of features are depicted by Tables 77, 78, 79, 80, and 81. Note that the values in the “Train”, “Test_1”, “Test_2”, and “Test_mean” columns regard the F1-score.

Table 77 – Sentiment group preprocessing result

Architecture	Stopped in Epoch	Train	Test_1	Test_2	Test_mean
1	246	0.71	0.69	0.69	0.69
2 (chosen)	446	0.72	0.69	0.705	0.6975
3	383	0.71	0.69	0.7	0.695
4	241	0.7	0.68	0.695	0.6875
			Final Mean Value		0.6925

Table 78 – Game Path group preprocessing result

Considered Positions	Architecture	Stopped in Epoch	Train	Test 1	Test 2	Test mean
1	1	253	0.695	0.68	0.685	0.6825
	2	295	0.69	0.67	0.69	0.68
	3	309	0.7	0.685	0.69	0.6875
	4	259	0.695	0.685	0.69	0.6875
				Final Mean Value		0.684375
2	1	250	0.69	0.67	0.68	0.675
	2	261	0.685	0.67	0.685	0.6775
	3	267	0.695	0.67	0.695	0.6825
	4	239	0.7	0.685	0.685	0.685
				Final Mean Value		0.68
3	1	235	0.685	0.655	0.68	0.6675
	2	266	0.69	0.67	0.685	0.6775
	3	262	0.7	0.67	0.685	0.6775
	4	247	0.695	0.67	0.68	0.675
				Final Mean Value		0.674375
4	1	244	0.69	0.665	0.685	0.675
	2	252	0.675	0.655	0.685	0.67

To be continued

Considered Positions	**Architecture**	**Stopped in Epoch**	**Train**	**Test 1**	**Test 2**	**Test mean**
	3	239	0.68	0.65	0.675	0.6625
	4	238	0.685	0.66	0.675	0.6675
				Final Mean Value		0.66875
5	1	240	0.685	0.665	0.68	0.6725
	2	232	0.685	0.665	0.685	0.675
	3	241	0.68	0.645	0.67	0.6575
	4	241	0.7	0.68	0.68	0.68
				Final Mean Value		0.67125
6	1	245	0.69	0.66	0.675	0.6675
	2	222	0.685	0.665	0.68	0.6725
	3	230	0.685	0.66	0.67	0.665
	4	232	0.685	0.66	0.68	0.67
				Final Mean Value		0.66875
7	1	236	0.69	0.66	0.68	0.67
	2	241	0.69	0.66	0.675	0.6675
	3	255	0.7	0.675	0.685	0.68
	4	232	0.68	0.65	0.67	0.66
				Final Mean Value		0.669375
8	1	259	0.695	0.67	0.685	0.6775
	2	246	0.7	0.67	0.685	0.6775
	3	241	0.695	0.665	0.68	0.6725
	4	233	0.695	0.67	0.675	0.6725
				Final Mean Value		0.675
9	1	235	0.68	0.655	0.68	0.6675
	2	231	0.68	0.645	0.67	0.6575
	3	249	0.7	0.67	0.68	0.675
	4	230	0.7	0.675	0.68	0.6775
				Final Mean Value		0.669375

To be continued

Considered Positions	Architecture	Stopped in Epoch	Train	Test 1	Test 2	Test mean
10	1	230	0.67	0.65	0.65	0.65
	2	234	0.7	0.67	0.675	0.6725
	3	232	0.69	0.665	0.675	0.67
	4	228	0.69	0.66	0.685	0.6725
				Final Mean Value		0.66625
11	1	237	0.685	0.665	0.635	0.65
	2	238	0.69	0.665	0.67	0.6675
	3	225	0.695	0.665	0.67	0.6675
	4	227	0.69	0.655	0.665	0.66
				Final Mean Value		0.66125
12	1	252	0.705	0.675	0.675	0.675
	2	237	0.705	0.675	0.685	0.68
	3 (chosen)	248	0.71	0.67	0.68	0.675
	4	233	0.705	0.675	0.655	0.665
				Final Mean Value		0.67375
13	1	226	0.685	0.66	0.625	0.6425
	2	250	0.705	0.675	0.685	0.68
	3	264	0.705	0.675	0.67	0.6725
	4	247	0.71	0.675	0.635	0.655
				Final Mean Value		0.6625
14	1	231	0.7	0.675	0.685	0.68
	2	223	0.68	0.66	0.635	0.6475
	3	228	0.685	0.655	0.625	0.64
	4	222	0.695	0.66	0.6	0.63
				Final Mean Value		0.649375
15	1	240	0.695	0.67	0.68	0.675
	2	226	0.675	0.65	0.555	0.6025
	3	260	0.7	0.665	0.675	0.67

To be continued

Considered Positions	Architecture	Stopped in Epoch	Train	Test 1	Test 2	Test mean
	4	224	0.695	0.665	0.61	0.6375
				Final Mean Value		0.64625
16	1	224	0.68	0.655	0.64	0.6475
	2	219	0.69	0.66	0.59	0.625
	3	254	0.705	0.67	0.685	0.6775
	4	233	0.695	0.655	0.625	0.64
				Final Mean Value		0.6475
17	1	219	0.685	0.655	0.55	0.6025
	2	238	0.695	0.67	0.65	0.66
	3	227	0.69	0.66	0.645	0.6525
	4	236	0.69	0.65	0.56	0.605
				Final Mean Value		0.63
18	1	227	0.69	0.665	0.66	0.6625
	2	233	0.695	0.66	0.465	0.5625
	3	233	0.7	0.655	0.58	0.6175
	4	218	0.69	0.665	0.65	0.6575
				Final Mean Value		0.625
19	1	227	0.69	0.67	0.67	0.67
	2	233	0.69	0.66	0.675	0.6675
	3	237	0.695	0.66	0.675	0.6675
	4	234	0.69	0.65	0.63	0.64
				Final Mean Value		0.66125
20	1	231	0.685	0.655	0.52	0.5875
	2	235	0.705	0.675	0.665	0.67
	3	226	0.69	0.66	0.595	0.6275
	4	229	0.675	0.65	0.665	0.6575
				Final Mean Value		0.635625
21	1	240	0.7	0.67	0.66	0.665

To be continued

Considered Positions	**Architecture**	**Stopped in Epoch**	**Train**	**Test 1**	**Test 2**	**Test mean**
	2	253	0.705	0.675	0.685	0.68
	3	233	0.695	0.66	0.675	0.6675
	4	228	0.695	0.655	0.66	0.6575
				Final Mean Value		0.6675
22	1	222	0.695	0.665	0.305	0.485
	2	235	0.685	0.66	0.655	0.6575
	3	225	0.695	0.66	0.605	0.6325
	4	242	0.705	0.675	0.67	0.6725
				Final Mean Value		0.611875
23	1	239	0.705	0.675	0.675	0.675
	2	229	0.695	0.665	0.635	0.65
	3	236	0.685	0.655	0.305	0.48
	4	232	0.695	0.66	0.64	0.65
				Final Mean Value		0.61375
24	1	237	0.7	0.68	0.635	0.6575
	2	234	0.695	0.67	0.68	0.675
	3	223	0.69	0.665	0.655	0.66
	4	221	0.695	0.67	0.615	0.6425
				Final Mean Value		0.65875
25	1	244	0.68	0.645	0.65	0.6475
	2	238	0.7	0.665	0.665	0.665
	3	230	0.69	0.66	0.655	0.6575
	4	221	0.69	0.66	0.525	0.5925
				Final Mean Value		0.640625
26	1	235	0.705	0.66	0.665	0.6625
	2	228	0.685	0.655	0.615	0.635
	3	230	0.695	0.66	0.52	0.59
	4	235	0.71	0.685	0.675	0.68

To be continued

Considered Positions	**Architecture**	**Stopped in Epoch**	**Train**	**Test 1**	**Test 2**	**Test mean**
				Final Mean Value		0.641875
27	1	233	0.69	0.66	0.55	0.605
	2	223	0.705	0.675	0.685	0.68
	3	228	0.685	0.645	0.495	0.57
	4	224	0.695	0.67	0.66	0.665
				Final Mean Value		0.63
28	1	231	0.7	0.68	0.64	0.66
	2	221	0.7	0.66	0.65	0.655
	3	244	0.695	0.655	0.67	0.6625
	4	230	0.695	0.665	0.645	0.655
				Final Mean Value		0.658125
29	1	228	0.68	0.655	0.515	0.585
	2	230	0.705	0.675	0.66	0.6675
	3	240	0.705	0.665	0.67	0.6675
	4	238	0.705	0.68	0.68	0.68
				Final Mean Value		0.65
30	1	263	0.705	0.67	0.68	0.675
	2	219	0.69	0.67	0.515	0.5925
	3	232	0.7	0.66	0.63	0.645
	4	233	0.7	0.67	0.67	0.67
				Final Mean Value		0.645625
31	1	227	0.69	0.655	0.685	0.67
	2	242	0.695	0.675	0.675	0.675
	3	236	0.705	0.67	0.675	0.6725
	4	227	0.7	0.665	0.65	0.6575
				Final Mean Value		0.66875
32	1	234	0.69	0.665	0.655	0.66
	2	230	0.695	0.67	0.5	0.585

To be continued

Considered Positions	Architecture	Stopped in Epoch	Train	Test 1	Test 2	Test mean
	3	250	0.705	0.66	0.685	0.6725
	4	227	0.695	0.66	0.64	0.65
				Final Mean Value		0.641875
33	1	224	0.7	0.67	0.315	0.4925
	2	233	0.705	0.67	0.68	0.675
	3	220	0.685	0.655	0.51	0.5825
	4	223	0.705	0.68	0.65	0.665
				Final Mean Value		0.60375

Table 79 – Macro Personality group preprocessing result

Architecture	Stopped in Epoch	Train	Test_1	Test_2	Test_mean
1	225	0.68	0.66	0.69	0.675
2	239	0.685	0.675	0.7	0.6875
3	264	0.685	0.675	0.695	0.685
4 (chosen)	347	0.7	0.685	0.695	0.69
			Final Mean Value		0.684375

Table 80 – Similarity group preprocessing result

Architecture	Stopped in Epoch	Train	Test_1	Test_2	Test_mean
1	283	0.69	0.695	0.68	0.6875
2	262	0.685	0.675	0.695	0.685
3	259	0.69	0.69	0.665	0.6775
4 (chosen)	299	0.69	0.69	0.695	0.6925
			Final Mean Value		0.685625

Table 81 – All features group preprocessing result

Architecture	Stopped in Epoch	Train	Test_1	Test_2	Test_mean
1 (chosen)	258	0.725	0.69	0.7	0.695
2	233	0.715	0.685	0.69	0.6875
3	222	0.71	0.675	0.685	0.68
4	230	0.71	0.685	0.685	0.685
			Final Mean Value		0.686875

For each group of features, churn labels were generated following the identified best architecture. Next, the results for a set of combinations between the psychological churn labels and the baseline's raw features are presented in Tables 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, and 103. In addition, Table 104 shows the ranking of all "Final Mean Values".

Table 82 – Raw features + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.42%	0.748	0.347	0.618	0.4825
RepTree	75.52%	0.748	0.73	0.711	0.7205
MLP	71.80%	0.714	0.585	0.67	0.6275
SVM	76.60%	0.755	0.729	0.727	0.728
			Final Mean Value		0.639625

Table 83 – Raw features + Macro Personality

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.40%	0.751	0.74	0.716	0.728
RepTree	75.25%	0.746	0.723	0.677	0.7
MLP	71.02%	0.707	0.653	0.649	0.651
SVM	76.90%	0.757	0.734	0.725	0.7295
			Final Mean Value		0.702125

Table 84 – Raw features + Macro Personality + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.75%	0.751	0.35	0.642	0.496
RepTree	75.92%	0.752	0.728	0.703	0.7155
MLP	72.33%	0.719	0.576	0.646	0.611
SVM	76.75%	0.757	0.727	0.725	0.726
			Final Mean Value		0.637125

Table 85 – Macro Personality + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	74.90%	0.755	0.737	0.745	0.741
RepTree	74.90%	0.755	0.737	0.745	0.741
MLP	74.30%	0.745	0.737	0.745	0.741
SVM	73.07%	0.74	0.726	0.741	0.7335
			Final Mean Value		0.739125

Table 86 – Raw features + Similarity

	Training		**Test F1-Score**		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.50%	0.752	0.628	0.643	0.6355
RepTree	75.87%	0.752	0.74	0.698	0.719
MLP	71.40%	0.71	0.596	0.666	0.631
SVM	77.07%	0.756	0.724	0.724	0.724
			Final Mean Value		0.677375

Table 87 – Raw features + Similarity + Macro Personality + Sentiment

	Training		**Test F1-Score**		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.63%	0.75	0.35	0.643	0.4965
RepTree	75.93%	0.752	0.739	0.739	0.739
MLP	70.85%	0.705	0.65	0.666	0.658
SVM	76.88%	0.758	0.727	0.729	0.728
			Final Mean Value		0.655375

Table 88 – Raw features + Similarity + Macro Personality

	Training		**Test F1-Score**		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.83%	0.755	0.74	0.716	0.728
RepTree	75.25%	0.746	0.723	0.677	0.7
MLP	71.45%	0.712	0.676	0.667	0.6715
SVM	76.93%	0.757	0.734	0.726	0.73
			Final Mean Value		0.707375

Table 89 – Raw features + Similarity + Sentiment

	Training		**Test F1-Score**		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.48%	0.749	0.343	0.644	0.4935
RepTree	75.53%	0.748	0.73	0.711	0.7205
MLP	71.88%	0.716	0.665	0.65	0.6575
SVM	76.63%	0.755	0.729	0.722	0.7255
			Final Mean Value		0.64925

Table 90 – Similarity + Macro Personality + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	74.90%	0.753	0.74	0.743	0.7415
RepTree	74.93%	0.754	0.737	0.745	0.741
MLP	74.33%	0.745	0.74	0.743	0.7415
SVM	72.73%	0.737	0.726	0.741	0.7335
			Final Mean Value		0.739375

Table 91 – Raw features + Game Path

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.78%	0.755	0.717	0.736	0.7265
RepTree	75.63%	0.751	0.72	0.727	0.7235
MLP	71.38%	0.71	0.676	0.66	0.668
SVM	77.13%	0.757	0.724	0.726	0.725
			Final Mean Value		0.71075

Table 92 – Raw features + Game Path + Similarity + Macro Personality + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.75%	0.751	0.733	0.727	0.73
RepTree	75.98%	0.753	0.739	0.739	0.739
MLP	71.43%	0.71	0.593	0.672	0.6325
SVM	76.85%	0.758	0.728	0.726	0.727
			Final Mean Value		0.707125

Table 93 – Raw features + Game Path + Macro Personality

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.53%	0.752	0.726	0.662	0.694
RepTree	75.13%	0.745	0.723	0.677	0.7
MLP	70.95%	0.706	0.651	0.684	0.6675
SVM	76.75%	0.756	0.729	0.725	0.727
			Final Mean Value		0.697125

Table 94 – Raw features + Game Path + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.53%	0.748	0.731	0.702	0.7165
RepTree	75.58%	0.749	0.73	0.711	0.7205
MLP	71.55%	0.713	0.629	0.669	0.649
SVM	76.68%	0.756	0.701	0.726	0.7135
			Final Mean Value		0.699875

Table 95 – Raw features + Game Path + Similarity

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.65%	0.756	0.661	0.66	0.6605
RepTree	75.48%	0.75	0.74	0.698	0.719
MLP	71.85%	0.714	0.65	0.644	0.647
SVM	77.13%	0.757	0.726	0.728	0.727
			Final Mean Value		0.688375

Table 96 – Game Path + Similarity + Macro Personality + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.00%	0.753	0.738	0.744	0.741
RepTree	74.98%	0.754	0.738	0.744	0.741
MLP	74.05%	0.741	0.738	0.744	0.741
SVM	72.73%	0.737	0.726	0.741	0.7335
			Final Mean Value		0.739125

Table 97 – Raw features + All features

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.63%	0.751	0.729	0.723	0.726
RepTree	75.80%	0.75	0.728	0.698	0.713
MLP	71.15%	0.709	0.567	0.654	0.6105
SVM	77.13%	0.76	0.727	0.722	0.7245
			Final Mean Value		0.6935

Table 98 – Raw features + All features + Game Path + Similarity + Macro Personality + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.75%	0.751	0.661	0.678	0.6695
RepTree	75.75%	0.751	0.739	0.739	0.739
MLP	71.48%	0.712	0.665	0.664	0.6645
SVM	76.93%	0.759	0.727	0.727	0.727
			Final Mean Value		0.7

Table 99 – Raw features + All features + Macro Personality

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.83%	0.752	0.729	0.722	0.7255
RepTree	75.45%	0.747	0.723	0.677	0.7
MLP	71.08%	0.706	0.626	0.668	0.647
SVM	76.90%	0.758	0.729	0.728	0.7285
			Final Mean Value		0.70025

Table 100 – Raw features + All features + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.70%	0.752	0.733	0.732	0.7325
RepTree	75.90%	0.752	0.73	0.711	0.7205
MLP	71.58%	0.712	0.579	0.647	0.613
SVM	77.08%	0.76	0.73	0.725	0.7275
			Final Mean Value		0.698375

Table 101 – Raw features + All features + Similarity

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.55%	0.751	0.729	0.723	0.726
RepTree	75.80%	0.75	0.728	0.698	0.713
MLP	71.93%	0.715	0.608	0.681	0.6445
SVM	77.18%	0.761	0.727	0.723	0.725
			Final Mean Value		0.702125

Table 102 – Raw features + All features + Game Path

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.55%	0.751	0.671	0.641	0.656
RepTree	75.85%	0.751	0.728	0.698	0.713
MLP	71.23%	0.709	0.639	0.642	0.6405
SVM	76.78%	0.757	0.725	0.724	0.7245
			Final Mean Value		0.6835

Table 103 – All features + Game Path + Similarity + Macro Personality + Sentiment

	Training		Test F1-Score		
Algorithm	**Accuracy**	**F1-Score**	**Test_1**	**Test_2**	**Test_mean**
C4.5	75.60%	0.759	0.741	0.746	0.7435
RepTree	75.68%	0.759	0.741	0.746	0.7435
MLP	74.80%	0.747	0.739	0.746	0.7425
SVM	73.05%	0.74	0.727	0.713	0.72
			Final Mean Value		0.737375

Table 104 – Ranking of churn prediction performance

Position	**Experiment**	**Final Mean Value**
1st	Sentiment + Macro Personality + Similarity + Game Paths	0.739
1st	Sentiment + Macro Personality + Similarity	0.739
1st	Sentiment + Macro Personality	0.739
4th	Sentiment + Macro Personality + Similarity + Game Paths + All	0.737
5th	Raw + Game Paths	0.71
6th	Raw + Macro Personality + Similarity	0.707
6th	Raw + Sentiment + Macro Personality + Similarity + Game Paths	0.707

To be continued

Position	Experiment	Final Mean Value
8th	**Baseline**	**0.706**
9th	Raw + Macro Personality	0.702
9th	Raw + Similarity + All	0.702
11th	Raw + Macro Personality + All	0.7
11th	Raw + Sentiment + Macro Personality + Similarity + Game Paths + All	0.7
13th	Raw + Sentiment + Game Paths	0.699
14th	Raw + Sentiment + All	0.698
15th	Raw + Macro Personality + Game Path	0.697
16th	Raw + All	0.693
17th	Raw + Similarity + Game Path	0.688
18th	Raw + Game Paths + All	0.683
19th	Raw + Similarity	0.677
20th	Raw + Sentiment + Macro Personality + Similarity	0.655
21st	Raw + Sentiment + Similarity	0.649
22nd	Raw + Sentiment + Macro Personality	0.637
23rd	Raw + Sentiment	0.63

As presented in Table 104, it is possible to notice that the baseline of 0.706 was overcome on some experiments (by seven different compositions of metrics). As we can see, the joining of the raw features with the psychological ones is not worth it, as the obtained F1-scores were higher than the baseline in only three cases of a total of 18. However, when only the psychological features are considered, the results overcome the baseline. It means that joining the tendency raw metrics perspective with the psychological one that contemplates short, mid, and long-term psychological aspects does not fit, being better to use only the psychological perspective instead of joining them.

The unification of all the four groups of psychological features to generate a unique churn label (i.e., the “All features” churn label) did not achieve the best result, showing that it is better to consider the features segregated than joined. It means that using the distinct notions of churn separately is better in this context than using a single churn view built based on all the psychological features. Also, as shown by Tables 77, 78, 79, 80, and 81, none psychological feature overcame the baseline alone, highlighting that they must be considered together as they are complementary to each other.

Accounting to the fact that the psychological features presented results that overcame the baseline, they are assessed as accurate. This means that the proposed method that converts actions/events into psychological metrics makes sense in the context of churn prediction of players in entertainment games.

Applying these findings to a real world scenario in a game company, the adopted approach would be the one that considers only the Sentiment, Macro Personality, Similarity, and Game Paths features. Even though the chosen approach presented the best final result equally to other two, this is the one with the highest number of psychological features, which encompasses more descriptions from a psychological perspective of what is a player churn. As these four metrics model the humans' psychological essence encompassing short, mid, and long-term aspects, the generated hypothesis by a classifier is expected to be resilient to future data for the same game as human nature is not expected to change regularly (LANKVELD, 2013; CARVER; SCHEIER, 2012). It means that the more diversified the psychological essences present on the training data are, the more resilient the model tends to be with unseen data. An interesting fact about these features is that they measure psychological essences that can be applied in other contexts besides gaming since emotions, sentiments, and personalities are universal influencers of human behavior.

Regardless of the resilience aspect of a model, the churn itself presents more challenges besides measuring the players' engagement (i.e., the identification of churners or non-churners). In this sense, even though a player presents psychological essences that means engagement according to the proposed metrics, the greatest "villain" that can make this kind of player churn is the amount of available content to be consumed. In situations where a player already achieved all the game challenges, and there are fewer social interactions than before, eventually, very engaged players tends to churn (ZHU; LI; ZHAO, 2010). Therefore, a good churn management must encompass a trustable churn prediction together with the release of new game contents.

Another interesting fact is that all the 119 raw metrics of the baseline can be replaced by only four metrics of the psychological perspective, which present better results. It foments that a good preprocessing is worth, as the generated features allow the classifiers to divide the hyperplane better. As the four preprocessed psychological features are just labels of "Yes" or "No" regarding different psychological perspectives, it is easier for a decision-maker to understand the rules identified by the classifier (transparent-box). For instance, if a decision tree was adopted, it is possible to check the more relevant feature that causes churn by looking at the top node in the tree (e.g., sentiment churn label). With this, counter-measures can be applied to develop new contents that change the players' sentiments to more desired levels, decreasing the churn rate (such as depicted in the final part of Subsection 5.3.2).

Besides the "available content villain", there is a "hidden villain" that cannot be mitigated by the game company, the players' life routine. Work, study, family, socialization, diseases, and cultural aspects can affect the players' life in a manner that they cannot play anymore (Burnout) (COOK, 2007). Given it, it is expected that the churn prediction on entertainment games present some distance to a perfect prediction. By looking at the best result obtained in this thesis, which can be seen as the state-of-the-art (the new baseline)[8], a mean F1-score of 0.739 presents a gap of 0.261 to be considered a perfect prediction. This gap represents a space of 26.1%, which can contain the effect of the aforementioned "villains", as well as not modeled players' behaviors. We expect that new researches improve the understanding of the players' behaviors present on this gap of 26.1%, decreasing it by increasing even more the knowledge about what makes a player plays and what does not. As a starting point, the unified human-being model has 60 GCs not approached by this thesis that would support better comprehension. Also, the unified players' model can provide additional knowledge with its 80 GCs.

Turning the focus to the imbalanceness of the dataset, such as depicted in Subsection 7.2.3, it is possible to notice that even though a class weight was adopted to mitigate it, the features showed a greater agreement on the non-churn perspective than on the churner one (the less present class), however with different hypotheses (as previously shown by the complementary aspects of them). These hypotheses divergence can be measured by the percentage in which each churn label differs to the others, such as shown by the column "Mean diff" in Tables 105, 106, and 107, regarding the subsets Training, Test_1, and Test_2, respectively. Another observation regards the best models mean accuracy, which was 74% (Tables 85, 90, and 96), following the same psychological features agreement of 74% (all churn + all non-churn of the Training subset, 29.57% and 45.15%, respectively), such as shown in Table 108.

Regarding the Training, the Sentiment, Macro Personality, Similarity, and All features churn labels presented at least a divergence of 10% to the other psychological features churn labels, being the Game Path label the one with the highest divergence (as depicted by Table 105). By looking at a general discordance between all churn labels in Table 108, it happened in 25.28% of the cases, highlighting the spread point of view of each feature. Note that this discordance also exists in the Test_1 and Test_2 with 20.17% and 38%, respectively.

[8] Even though the actual churn labels for the test subsets were needed in the preprocessing (what were not available during the Data Mining competition), we understand that the comparison to the baseline keeps fair, as these labels do not influence the definition of the Training classifier. However, we also understand that special care should be taken to avoid dishonest behaviors during a competition. Thus, to keep the same secrecy of the test subsets labels as adopted by this thesis, a competition should provide means to the competitors to apply feature engineering techniques to the test subsets encompassing the test actual churn labels without turning them public. An option would be a server that receives a set of features, a classifier and adds the actual labels, returning only the predicted values. Note that this kind of approach keeps the test labels hidden, not allowing any dishonest behavior, such as happened in this thesis.

Table 105 – Churn labels divergence between the psychological features in the Training subset

	To					
From	**Sentiment**	**Macro Personality**	**Similarity**	**Game Paths**	**All features**	**Mean diff**
Sentiment	-	11.25%	10.58%	15.55%	10.05%	11.86%
Macro Personality	11.25%	-	10.50%	14.88%	11.53%	12.04%
Similarity	10.58%	10.50%	-	12.00%	13.50%	11.65%
Game Paths	15.55%	14.88%	12.00%	-	14.23%	14.17%
All features	10.05%	11.53%	13.50%	14.23%	-	12.33%

Table 106 – Churn labels divergence between the psychological features in the Test_1 subset

	To					
From	**Sentiment**	**Macro Personality**	**Similarity**	**Game Paths**	**All features**	**Mean diff**
Sentiment	-	7.13%	6.20%	12.67%	9.00%	8.75%
Macro Personality	7.13%	-	7.50%	12.70%	9.03%	9.09%
Similarity	6.20%	7.50%	-	12.83%	10.50%	9.26%
Game Paths	12.67%	12.70%	12.83%	-	11.30%	12.38%
All features	9.00%	9.03%	10.50%	11.30%	-	9.96%

Table 107 – Churn labels divergence between the psychological features in the Test_2 subset

	To					
From	**Sentiment**	**Macro Personality**	**Similarity**	**Game Paths**	**All features**	**Mean diff**
Sentiment	-	10.70%	11.30%	25.70%	15.30%	15.75%
Macro Personality	10.70%	-	8.60%	26.20%	16.60%	15.53%
Similarity	11.30%	8.60%	-	28.07%	20.60%	17.14%
Game Paths	25.70%	26.20%	28.07%	-	20.33%	25.08%
All features	15.30%	16.60%	20.60%	20.33%	-	18.21%

The Game Paths presented 14.17%, 12.38%, and 25.08% of divergence (bias) on subsets Training, Test_1, and Test_2, respectively. By contrast, the lower disagreements per subset regarded the Similarity, Sentiment, and Macro Personality with values of 11.65%, 8.75%, and 15,53% on subsets Training, Test_1, and Test_2, respectively. This Game Path bias regards two significant situations: (1) exclusive non-churner and churner segments on the Training subset do not carry the same notion on the tests subsets (as shown by Table 109), and (2) the Positions range in the Test_2 subset is more extensive (115 possible Positions) than the Training subset (33 possible Positions). The bias between the Training

Table 108 – All churn labels agreements and disagreements

Subset	**All churn**	**All non-churn**	**Diffs**
Training	29.57%	45.15%	25.28%
Test_1	31.10%	48.73%	20.17%
Test_2	24.30%	37.70%	38.00%

and Test_1 subsets is lower since they share a similar number of possible Positions (33 for Training and 35 for Test_1).

Table 109 – Divergence between churn notions in the Game Path Segments

Subsets	**Segments with the same churn notion**	**Segments with the opposite churn notion**	**Segments with mixed churn notion**
Training to Test_1	19.44%	8.33%	72.23%
Training to Test_2	13.88%	8.33%	77.79%

For an ideal Game Path analysis with a decreased bias, the Game Paths should be computed considering all the players' actions since their first gameplay. However, it is impossible in the considered Blade&Soul dataset, such as previously pointed out in the final part of Subsection 6.3.2.3. Despite that, interestingly, the addition of the Game Path churn label to the raw data presented a better result compared to the baseline, but lower than the best results. It means that even though the Game Path has a bias, its induced hypothesis fits somehow to the notion present on the raw features.

7.4 Psychological Data Enhancement

Given the positive assessment of the proposed method's resultant psychological profile, it is proposed the term "Psychological Data Enhancement" as a reference point to the processing technique that converts players' actions into psychological features presented in this thesis. Therefore, the term "Psychological Data Enhancement" is coined as:

Psychological Data Enhancement

The process of translating historical actions and events that affect people into short, mid, and long-term psychological features.

8 Conclusions and Future Works

This work was motivated by the identified gap in the Game Analytics field that regards the identification of players' psychological profile based on usage data according to multiple psychological models, and in view of it, the problem statement "Is it possible to identify psychological profiles encompassing multiple psychological models on digital games based on usage data?" was proposed. To understand the subjects around this topic, 11 research questions were suggested and answered through an SLR, the Unification Explorer Framework (UEF), and the method proposition. Moreover, all the findings regarding the psychological aspects outside of the game context were validated with specialists. Also, the generated psychological profile was assessed in the churn context. Summing all the findings of this thesis, the answer to the problem statement is "Yes, it is possible by following four main steps: (1) the identification of psychological models applied to games, (2) the unification of them following some heuristic, (3) the extraction of the unified model characteristics from data to provide a psychological profile, and (4) the assessment of the resultant profile". In this thesis, the first step was accomplished by the proposed SLR that identified 46 players' models and 21 human-being's models; the second by the UEF proposition and application, where the adopted heuristic regarded a quantitative perspective linked to the holism concept; the third by the method proposition that generates psychological metrics related to short, mid, and long-terms aspects; and the fourth by the assessment of specialists, the literature support, and the churn prediction improvement.

A summarized answer for each RQ can be seen next:

- **RQ1:** *What is a psychological aspect?* A psychological aspect can be understood in more specific terms such as affect, emotion, sentiment, personality, personality traits, competence, and human needs.
- **RQ2:** *What is a psychological profile?* In our concept, a psychological profile is a set of characteristics (i.e., psychological aspects or features) that can describe one's behavior or way of being.
- **RQ3:** *What are the psychological models applied to games?* So far, the 67 models identified in the proposed SLR.
- **RQ4:** *Is it possible to link a profile of one model to the profile of another model?* Yes, it is. The UEF's "Joining step" exemplifies it.
- **RQ5:** *Can psychological models be ranked?* From a quantitative perspective, models can be ranked following the UEF's "Ranking step".

- **RQ6:** *Is it possible to combine models?* Yes, it is. When the UEF proposes a new unified model, it is proposing a model that is the result of models combinations.
- **RQ7:** *Is there a general psychological model that can portray all the players' aspects?* Considering the set of identified models and excluding the models generated by the UEF, no, there is not.
- **RQ8:** *Are all models applicable to all game genres?* Excluding a hypothetical case where a unified model was built based on at least one psychological model from each possible genre, no, they are not.
- **RQ9:** *What are the advantages and disadvantages of using psychological models?* The UEF presented a set of pros and cons of using psychological models, concluding that their adoption is worth it, given that the more descriptions are provided to comprehend the players, the better. Also, this better comprehension was exemplified as useful in the following context: the mitigation of risky situations (e.g., churn), the development of more assertive content (e.g., through community management), and the propositions of more believable NPCs.
- **RQ10:** *To what extent characteristics of usage data can be used to identify psychological profiles?* As far as our efforts allowed us, to identify human needs, emotions, sentiments, and personality traits.
- **RQ11:** *How an identified profile on usage data can be assessed?* Besides the assessments performed by psychologists, the adoption of psychological features on the prediction of risk situations (e.g., churn) can provide insights about the support or not of these features give a better performance (following an evaluation metric) or not, assessing them as accurate or not.

The proposed method was designed to identify as many as possible psychological aspects on usage data. So far, human needs, emotions, sentiments, and personality traits were theoretically identified. The method was applied to the Blade&Soul game, an MMORPG that contemplates all the method's essential assumptions. However, unfortunately, this game dataset does not provide the required information to compute the social emotions. Nevertheless, the internal emotions could provide insightful findings, as they were used to generate players' sentiments that portray, in an individualized manner, what are the players' feelings in achieving their human needs.

The resultant psychological profile is a combination of what players desire (i.e., their personalities) and their success or not in attaining such wishes (i.e., their sentiments). Moreover, this profile was proposed based on an interesting characteristic identified in the general human-being model, the "Aggregated psychological essences" property. With this, it was possible to propose short, mid, and long-term psychological metrics based only on what happens to a person (i.e., the historical sequence of actions and events that affect him/her).

The proposed profile supports better management of a game usage lifecycle from different ways, such as in assessing what players' desires, their approval over new game upgrades, their success rate in achieving their needs, the identification of an ideal challenge degree, the churn management from a reactive or proactive perspective, the proposition of new contents, and the measurement of content consumption. Also, the proposed profile overcame the state-of-the-art baseline regarding the churn prediction. These benefits regard the simple fact of better understanding players, where the more aspects are comprehended, the better.

To present this thesis benefits from a game design or game producer point of view, the questions presented in Chapter 1 are now answered based on the applicability of the proposed psychological metrics.

- *What are the individual players' motivations for a given game?* The individual players' motivations to play a specific game can be identified based on the Macro Spectrum metrics, as they portray what players desire or not to chase considering the game content available to them, also encompassing the notion of players' priorities.
- *What is the degree of similarity between the active players' behavior in a game?* This kind of similarity can be identified considering two perspectives, the Macro and Micro Spectrums. The Macro Spectrum similarity metric gives the idea of how similarly players chase the same sets of human needs, whereas the Micro Spectrum one provides a measure of how similarly or not players choose to do the same sequence of choices.
- *Is a game entertaining its players with a comfortable/desirable degree of challenge?* This kind of identification can be achieved by looking at two kinds of metrics at the same time, the players' sentiments and the churn rate. On the one hand, if the churn rate is low and the sentiments are positive, it means that players are enjoying the difficult degree that allows them to have more successes than failures regarding a given human need attainment (the idea of achievable challenges); however, if the sentiments were negative, it means that players enjoy a more difficult game, where it is harder to attain their needs. On the other hand, if the churn rate is high and the player sentiments are positive, it means that the game is too easy and the players do not like that; besides, if the players' sentiments were negative, it means that the game is too difficult to them (unpleasurable). This kind of analysis is very attached to the Flow concept (DECI; RYAN, 1985; DECI; RYAN, 1995; HUIZINGA, 2014).

- *Was a given game upgrade successful? What was a possible cause?* This identification can be performed by comparing two snapshots of the players' sentiments, one before and the other after the game upgrade. If the snapshots present change on its colors (green for positive and red for negative sentiments) and it is attached to the increase of the churn rate, it is possible to say that the upgrade was not successful and point a possible cause (e.g., a sentiment that changed from positive to negative). Also, if the churn rate decreases, the upgrade can be seen as successful and its cause as a change in the players' sentiments (e.g., from negative to positive). This kind of analysis allows a validation over changes on specific game mechanisms; for example, if a new combat system was incorporated in the game and the Power need (that is linked to this kind of mechanism) changed its color attached to an increase in the churn rate, it is possible to assume that this new mechanism was not approved by the players. The same analysis can be performed to any human need, such as assessing if the players liked the new kinds of social interactions (linked to the Affiliation need), the new reforging system (linked to the Achievement need), the new opportunities to gathering items (linked to the Materialism need), the new quests to be completed (linked to the Information need), or the new possibility to get marriage in-game (linked to the Sensual need).

- *When should I release a game upgrade?* The release of a game upgrade must consider several aspects, such as the amount of available content to be consumed by players, seasonality, competitors' games, and the availability of new contents to be released. The proposed profile has a metric that can support the first aspect, the Available Content metric. This metric is based on the Game Paths idea and measures the amount of content available to be consumed by players in the same way as other players did. A special nature of this metric is that it is not attached to the game design but by the players' choices. It means that this metric can identify inside the whole game design content what players prefer to consume and give a measure of how much is yet available in this perspective. The motivation behind this metric is that different players playing the same game can have different motivations to do it, meaning that some parts of the game content may not be of interest to different players. Therefore, it is essential to measure this available content from a personality perspective, as done by this metric.

- *What are the game design components that players most chase in a given game?* By considering a game as a source of attainment of human needs, the idea of design components can be comprehended then as candidate tools available to the player that allows them to attain their human needs. If we link to each design component the human needs affected by it and checking the Macro Spectrum ranking, it is possible to verify what kinds of game design components are most chased by players. Game designers can perform this linkage between human needs and design components by looking at the general human-being model's GCs descriptions linked to the human needs (PsyHM1, PsyHM4, PsyHM5, PsyHM6, PsyHM7, and PsyHM8) and checking for each design component which human need is related to these descriptions. For example, social interactions are linked to the Affiliation need, the arrangements of objects to the Materialism need, the capacity to attack others to the Power need, recognition to the Achievement need, exploration to the Information need, and sensuous expressions with the Sensual need.

- *What components of a game design should be added to increase engagement?* To do it, first, it is needed to identify what players desire, which can be given by the Macro Spectrum metrics. Bearing in mind that the game content consumption is always present while players are playing and that even engaged players can leave the game if he/she consumes all the desired content, new content should always be provided to players. If it was identified that players enjoy the Affiliation and Power needs, new design components should be added regarding these needs, as it is known that players enjoy them. Also, it is possible to offer new kinds of components that attain a human need not previously offered by the game content (e.g., the Sensual need). Moreover, even though it is known that the active players like some kinds of design components, it is essential to monitor the change in these players' behavior after the release to verify if the challenge degrees of these new components are pleasurable or not to them.

- *What components of a game design should be removed or modified to increase engagement?* The Macro Spectrum presents two perspectives, a ranking of human needs chase and the identification of chase or not of human needs. If a given game offers components related to a need not chased by players, these components can be removed without problems. For example, if a game has components regarding social interactions, but all players are unsocial (not chasing the Affiliation need), these components can be removed. Also, harmful components can be identified by looking at the sentiments of players. For example, if a given player abandoned a game due to dying consecutive times, which impaired its Power need until reaching a negative sentiment. If this impairment is something expected to happen in terms of game design, the linked design components should be revisited. One option could be the removal of this component in the game content to avoid players reaching the same state, or to modify it by reducing the degree of impairment, as it leads players to leave the game. This kind of analysis can be performed for each human need individually.

- *What is the earliest moment when I can identify churn candidates?* Considering the Game Path idea, churn candidates can be identified since the first action of players in-game, even though these players are very engaged in continuing playing.

- *What would be an appropriate way to understand individual churn behaviors?* By joining what players desire and their success rate in attaining their wishes, it is possible to identify when a player is negatively affected to continuing playing. For example, considering the Macro Spectrum metrics, it is possible to identify what players like to do in-game, and by looking at their sentiments while playing, it is possible to identify if they are successful or not in attaining their wishes. By providing these two perspectives to a classifier together with a churn label, it is possible to identify churn rules that carry psychological constructs. It is essential to highlight that these two perspectives should be analyzed together. Let us assume an unsocial player, if only the sentiment is observed (which means that we do not know if he/she is an unsocial player or not), and this player has a negative sentiment, a possible conclusion is that this player may not be engaged, but if the Macro Spectrum perspective is considered, it is possible to see that the Affiliation aspect to this player has no importance; thus the negative sentiment carries no meaning to him/her. By contrast, if he/she is a very active social player, this negative sentiment is an alert and a relevant piece of information to consider while prospecting this player churning probability. Note that this kind of analysis can be performed for each human need.

- *What makes players start or continue playing?* Basically, the opportunity to attain human needs. However, these abstract concepts of human needs can be understood in more specific terms by looking at the GCs present on the unified players' model, which portray players' motivations (like combat challenges), preferences (such as social interactions), status (as the degrees of mastery over the game mechanisms), and the general reasons to play games (as a pastime activity, a need or a desire). It means that while players can identify opportunities to attain their desired human needs, they will continue playing, as well as accepting to play new games.

The proposed profile also presents limitations, as its features are assumed as being quantitative, disregarding any qualitative aspect. For instance, there is no procedure to differentiate inside a number of actions, which is unique and entailed a special pleasure to the player. Also, the proposed method does not encompass any artistic aspect of games, as no procedure is provided to measure the players' pleasure linked to sounds and images. Even though this artistic limitation is, yet, an obscure subject to the Game Analytics field, the proposed method can mitigate the qualitative problem by the proposed joining to the Game Refinement Theory, where the findings of both, until parallel research fields, can be combined to support a better understanding of players. A final limitation regards the GCs that cannot be identified based only on usage data, such as the LtMotSour1 and LtMotSour2, which regards the source of a person's motivation as being internal or external. To identify such aspects, questionnaires should be needed (an invasive approach). As the proposed method is non-invasive (to cover a more significant number of players), some GCs cannot be approached.

In addition, the proposed method differs from the state-of-the-art in four points: (1) it was proposed based on multiple psychological models, where a criterion to define a model was adopted; (2) the method was applied to a real game dataset that contains a half of billion instances and not to an experimental game with few instances; (3) it was experimentally assessed in a churn prediction problem; and (4), the proposition of a term that references the proposed method main idea ("Psychological Data Enhancement").

Besides the resultant profile and methods contributions, the proposed SLR also provided exciting findings, as a list of psychological models applied to games was identified regarding two contexts, players' behaviors and general human behavior. In addition, the UEF application to these contexts resulted in two general models' propositions, where other researchers can use such models to try to identify other psychological aspects not approached by this thesis.

Regarding the technology transfer, a desktop system that contemplates all the method metrics were developed, the 3PIS. Moreover, the 3PIS was adapted to, theoretically, be applied to any game that contains telemetry data with players' actions.

As final remarks, 16 future works were suggested, being they:

1. To apply the UEF to the two general models, resulting in a proposition of a new and enhanced model or the promotion of one of them as a general model in this new context.

2. To apply another SLR to identify possible new models and perform another execution of the UEF.

3. To identify more GCs on usage data based on the identified general models.

4. To identify the social emotions on usage data.

5. To group game genres according to which types of human needs they attain.

6. To add the intensity idea in the actions (for example, consider the amount of money acquired in a "GetMoney" action; a qualitative approach).

7. To generate six new metrics regarding the GRV applied to each of the six human needs.

8. To identify unprecedented events on data (i.e., when a player faces a new challenge or accomplishes a challenge for the first time).

9. To identify affect and competencies in usage data.

10. To apply the proposed method to other games and other game genres apart from the MMORPG one.

11. To identify the ideal level of the available content metric to release a game upgrade.

12. To increment the Game Paths capabilities to allow their generations considering segregated perspectives for the same players.

13. To associate the generated psychological profile to other problems besides a boolean prediction of churn, like survival time prediction, monetization candidates identification (i.e., what are the players that could turn from free user to pay users), bot detection, and player simulation.

14. To retrieve the full potential of the Game Paths by computing them since the first gameplay of players and then check the proactive churn management potential.

15. To propose a map that suggests GCs from the proposed general models linked to a specific game genre, facilitating its usage.

16. To find a manner to weigh the GCs considering their complexity and potential benefits to the game design.

Bibliography

ALLEN, Bem P. *Personality theories: Development, growth, and diversity.* [S.l.]: Psychology Press, 2015. Citado na página 44.

ANANKIN, Sergei. *Your Inner Fish That Wants to Play: What Anthropology Has to Say about Games.* 2018. Disponível em: <http://www.gamasutra.com/blogs/SergeiAnankin/20181204/329082/Your_Inner_Fish_That_Wants_to_Play_What_Anthropology_Has_to_Say_about_Games.php>. Citado na página 89.

ARTHUR, David; VASSILVITSKII, Sergei. k-means++: The advantages of careful seeding. In: SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms.* [S.l.], 2007. p. 1027–1035. Citado na página 106.

AUNG, Myat; BONOMETTI, Valerio; DRACHEN, Anders; COWLING, Peter; KOKKINAKIS, Athanasios; WADE, Alex. Predicting skill learning in a large, longitudinal moba dataset. In: IEEE. *2018 IEEE Conference on Computational Intelligence and Games (CIG).* [S.l.], 2018. p. 117–124. Citado na página 27.

AUNG, Myat; DEMEDIUK, Simon; SUN, Yuan; TU, Ye; ANG, Yu; NEKKANTI, Siva; RAGHAV, Shantanu; KLABJAN, Diego; SIFA, Rafet; DRACHEN, Anders. The trails of just cause 2: spatio-temporal player profiling in open-world games. In: ACM. *Proceedings of the 14th International Conference on the Foundations of Digital Games.* [S.l.], 2019. p. 41. Citado na página 27.

AUYANG, Sunny Y. *Foundations of complex-system theories: in economics, evolutionary biology, and statistical physics.* [S.l.]: Cambridge University Press, 1998. Citado na página 33.

BARKER, Valerie. Older adolescents' motivations for social network site use: The influence of gender, group identity, and collective self-esteem. *Cyberpsychology & behavior*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801-5215 USA, v. 12, n. 2, p. 209–213, 2009. Citado na página 27.

BARTLE, Richard. Hearts, clubs, diamonds, spades: Players who suit muds. *Journal of MUD research*, v. 1, n. 1, p. 19, 1996. Citado 3 vezes nas páginas 32, 45, and 61.

BATEMAN, Chris; BOON, Richard. *21st Century Game Design. Hingham, MA: Charles River Media.* [S.l.]: Inc, 2006. Citado 6 vezes nas páginas 13, 45, 61, 81, 128, and 147.

BATES, Joseph et al. The role of emotion in believable agents. *Communications of the ACM*, Citeseer, v. 37, n. 7, p. 122–125, 1994. Citado na página 80.

BEN-ZE'EV, A. *The subtlety of Emotions (vol. 12).* [S.l.]: Cambridge, MA, USA: MIT Press, 2000. Citado 2 vezes nas páginas 42 and 170.

BERTENS, Paul; GUITART, Anna; CHEN, Pei Pei; PERIÁÑEZ, África. A machine-learning item recommendation system for video games. In: IEEE. *2018 IEEE Conference on Computational Intelligence and Games (CIG).* [S.l.], 2018. p. 385–388. Citado 2 vezes nas páginas 27 and 89.

BERTENS, Paul; GUITART, Anna; PERIÁÑEZ, África. Games and big data: A scalable multi-dimensional churn prediction model. In: IEEE. *2017 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2017. p. 33–36. Citado na página 26.

BLACOW, Glen. Aspects of adventure gaming. *Different worlds*, v. 10, 1980. Citado 2 vezes nas páginas 45 and 62.

BONOMETTI, Valerio; RINGER, Charles; HALL, Mark; WADE, Alex R; DRACHEN, Anders. Modelling early user-game interactions for joint estimation of survival time and churn probability. In: IEEE. *2019 IEEE Conference on Games (CoG)*. [S.l.], 2019. p. 1–8. Citado 2 vezes nas páginas 28 and 74.

BOSTAN, Barbaros. Player motivations: A psychological perspective. *Computers in Entertainment (CIE)*, ACM, v. 7, n. 2, p. 22, 2009. Citado 10 vezes nas páginas 76, 78, 79, 83, 160, 165, 205, 208, 209, and 223.

BOYLE, Elizabeth A; CONNOLLY, Thomas M; HAINEY, Thomas; BOYLE, James M. Engagement in digital entertainment games: A systematic review. *Computers in human behavior*, Elsevier, v. 28, n. 3, p. 771–780, 2012. Citado na página 42.

BRAYSHAW, Mike; GORDON, Neil. Using motivation derived from computer gaming in the context of computer based instruction. In: IEEE. *SAI Computing Conference (SAI), 2016*. [S.l.], 2016. p. 828–832. Citado 2 vezes nas páginas 45 and 61.

BULATOV, Vadim. *How Marketing drives engagement in free to play games*. 2018. Disponível em: <https://www.gamasutra.com/blogs/VadimBulatov/20180220/314756/How_Marketing_drives_engagement_in_free_to_playgames.php>. Citado 2 vezes nas páginas 45 and 61.

CAILLOIS, Roger. *Man, play, and games*. [S.l.]: University of Illinois Press, 2001. Citado 2 vezes nas páginas 45 and 61.

CALVO, Rafael A; D'MELLO, Sidney; GRATCH, Jonathan Matthew; KAPPAS, Arvid. *The Oxford handbook of affective computing*. [S.l.]: Oxford University Press, USA, 2015. Citado 2 vezes nas páginas 27 and 28.

CAMILLERI, Elizabeth; YANNAKAKIS, Georgios N; LIAPIS, Antonios. Towards general models of player affect. In: IEEE. *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. [S.l.], 2017. p. 333–339. Citado na página 74.

CAMMARATA, Aaron; KOSTER, Raph; Google's Advanced Technology and Projects (ATAP) group. *The Trust Spectrum*. 2018. Disponível em: <https://www.gamasutra.com/blogs/RaphKoster/20180316/315615/The_Trust_Spectrum.php>. Citado 2 vezes nas páginas 45 and 61.

CARPINETO, Claudio; ROMANO, Giovanni. *Concept data analysis: Theory and applications*. [S.l.]: John Wiley & Sons, 2004. Citado na página 109.

CARVER, Charles S; SCHEIER, Michael F. *Perspectives on personality*. [S.l.]: Pearson education, 2012. Citado 5 vezes nas páginas 41, 42, 171, 199, and 288.

CASTRO, Emiliano G; TSUZUKI, Marcos SG. Churn prediction in online games using players' login records: A frequency analysis approach. *IEEE Transactions on Computational Intelligence and AI in Games*, IEEE, v. 7, n. 3, p. 255–265, 2015. Citado na página 26.

CASVEAN, Tulia Maria. Understanding the videograme genre: a qualitative analysis of the "playing contract". *Romanian Journal of Communication and Public Relations*, v. 17, n. 3, p. 59–68, 2015. Citado na página 31.

CHARLTON, William et al. *Aristotle's Physics: books I and II*. [S.l.]: Oxford University Press, 1983. Citado na página 40.

CHEN, Pei Pei; GUITART, Anna; RÍO, Ana Fernández del; PERIÁÑEZ, África. Customer lifetime value in video games using deep learning and parametric models. In: IEEE. *2018 IEEE International Conference on Big Data (Big Data)*. [S.l.], 2018. p. 2134–2140. Citado na página 27.

CHETPRAYOON, Panumate. Various approaches to improving entertainment impact in games. 2016. Citado na página 94.

CHOU, Chien; TSAI, Meng-Jung. Gender differences in taiwan high school students' computer game playing. *Computers in Human Behavior*, Elsevier, v. 23, n. 1, p. 812–824, 2007. Citado 2 vezes nas páginas 45 and 62.

CLARK, Oscar. *Games as a service: How free to play design can make better games*. [S.l.]: Focal Press, 2014. Citado na página 24.

COLWELL, John. Needs met through computer game play among adolescents. *Personality and Individual Differences*, Elsevier, v. 43, n. 8, p. 2072–2082, 2007. Citado 3 vezes nas páginas 45, 61, and 128.

COOK, Daniel. *The Circle of Life: An Analysis of the Game Product Lifecycle*. 2007. Disponível em: <https://www.gamasutra.com/view/feature/129880/the_circle_of_life_an_analysis_of_.php>. Citado 11 vezes nas páginas 10, 24, 30, 31, 32, 45, 61, 80, 88, 91, and 289.

COWLEY, Benjamin Ultan. *Player profiling and modelling in computer and video games*. Tese (Doutorado) — University of Ulster, 2009. Citado na página 43.

CROSS, Katherine. *Opinion: World of Warcraft: Classic will disappoint you*. 2018. Disponível em: <http://www.gamasutra.com/view/news/331148/Opinion_World_of_Warcraft_Classic_will_disappoint_you.php>. Citado na página 88.

CZISIKSZENTMIHALYI, M. *Flow-The Psychology of Optimal Experience*. [S.l.]: Harper & Row, 1990. Citado 4 vezes nas páginas 24, 45, 72, and 205.

DECI, Edward; RYAN, Richard M. *Intrinsic motivation and self-determination in human behavior*. [S.l.]: Springer Science & Business Media, 1985. Citado 5 vezes nas páginas 45, 72, 174, 205, and 294.

DECI, Edward L; RYAN, Richard M. Human autonomy. In: *Efficacy, agency, and self-esteem*. [S.l.]: Springer, 1995. p. 31–49. Citado 5 vezes nas páginas 45, 72, 174, 205, and 294.

DEMETROVICS, Zsolt; URBÁN, Róbert; NAGYGYÖRGY, Katalin; FARKAS, Judit; ZILAHY, Dalma; MERVÓ, Barbara; REINDL, Antónia; ÁGOSTON, Csilla; KERTÉSZ, Andrea; HARMATH, Eszter. Why do you play? the development of the motives for online gaming questionnaire (mogq). *Behavior research methods*, Springer, v. 43, n. 3, p. 814–825, 2011. Citado 3 vezes nas páginas 45, 62, and 128.

DICKEY, Michele D. Game design and learning: A conjectural analysis of how massively multiple online role-playing games (mmorpgs) foster intrinsic motivation. *Educational Technology Research and Development*, Springer, v. 55, n. 3, p. 253–273, 2007. Citado 2 vezes nas páginas 45 and 62.

DRACHEN, Anders. *Behavioral Profiling: How to Build Useful Player Models*. 2018. Disponível em: <https://www.gamasutra.com/blogs/AndersDrachen/20180116/313099/Behavioral_Profiling_How_to_Build_Useful_Player_Models.php>. Citado na página 26.

DRACHEN, Anders; SCHUBERT, Matthias. Spatial game analytics. In: *Game Analytics*. [S.l.]: Springer, 2013. p. 365–402. Citado na página 27.

DRACHEN, Anders; SIFA, Rafet; BAUCKHAGE, Christian; THURAU, Christian. Guns, swords and data: Clustering of player behavior in computer games in the wild. In: IEEE. *2012 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2012. p. 163–170. Citado na página 26.

DRACHEN, Anders; THURAU, Christian; SIFA, Rafet; BAUCKHAGE, Christian. A comparison of methods for player clustering via behavioral telemetry. *arXiv preprint arXiv:1407.3950*, 2014. Citado na página 26.

ECKARDT, Barbara Von. *What is cognitive science?* [S.l.]: MIT press, 1995. Citado na página 145.

EGLESZ, Dénes; FEKETE, Istvan; KISS, Orhidea Edith; IZSO, Lajos. Computer games are fun? on professional games and players' motivations. *Educational Media International*, Taylor & Francis, v. 42, n. 2, p. 117–124, 2005. Citado 2 vezes nas páginas 45 and 62.

EKMAN, Paul. An argument for basic emotions. *Cognition & emotion*, Taylor & Francis, v. 6, n. 3-4, p. 169–200, 1992. Citado 4 vezes nas páginas 41, 45, 72, and 81.

EL-NASR, Magy Seif; DRACHEN, Anders; CANOSSA, Alessandro. *Game analytics*. [S.l.]: Springer, 2016. Citado 2 vezes nas páginas 26 and 43.

EYSENCK, HJ. The biological basis of personality. Spring-field, Ill., 1967. Citado 3 vezes nas páginas 45, 72, and 142.

EYSENCK, Hans Jurgen. The inequality of man. 1973. Citado 2 vezes nas páginas 45 and 72.

FAIRCLOUGH, Stephen H. Fundamentals of physiological computing. *Interacting with computers*, Oxford University Press Oxford, UK, v. 21, n. 1-2, p. 133–145, 2008. Citado 4 vezes nas páginas 24, 32, 45, and 61.

FERNANDES, Lucas V; CASTANHO, Carla D; JACOBI, Ricardo P. A survey on game analytics in massive multiplayer online games. In: IEEE. *2018 17th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. [S.l.], 2018. p. 21–2109. Citado na página 26.

FONG, Henry. *Want to Boost Your Game's ROI? Tap Into Your Community*. 2019. Disponível em: <https://www.gamasutra.com/blogs/HenryFong/20190530/343415/Want_to_Boost_Your_Games_ROI_Tap_Into_Your_Community.php>. Citado na página 150.

FRENCH, Vera V. The structure of sentiments: I. a restatement of the theory of sentiments. *Journal of Personality*, Wiley Online Library, v. 15, n. 1, p. 247–282, 1947. Citado na página 205.

FREUD, Sigmund. Recommendations to physicians practising psycho-analysis. Classics in Psychoanalytic Technique (1990), 1912. Citado na página 40.

FREUD, Sigmund. Beyond the pleasure principle (standard ed., vol. 18). *London: Hogarth*, 1920. Citado na página 40.

FRIJDA, Nico H. *The emotions*. [S.l.]: Cambridge University Press, 1986. Citado 3 vezes nas páginas 45, 72, and 81.

FROSTLING-HENNINGSSON, Maria. First-person shooter games as a way of connecting to people:"brothers in blood". *CyberPsychology & Behavior*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 12, n. 5, p. 557–562, 2009. Citado 3 vezes nas páginas 45, 62, and 128.

GARDA, Maria B. Neo-rogue and the essence of roguelikeness. *Homo Ludens*, Polskie Towarzystwo Badania Gier (Games Research Association of Poland), v. 1, n. 5, p. 59–72, 2013. Citado na página 31.

GOLDBERG, Lewis R. An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, American Psychological Association, v. 59, n. 6, p. 1216, 1990. Citado 4 vezes nas páginas 45, 72, 74, and 75.

GOLDREI, DC. *Classic Set Theory: For Guided Independent Study*. [S.l.]: Routledge, 2017. Citado na página 109.

GOLEMAN, Daniel. *Working with emotional intelligence*. [S.l.]: New York: Bantam Books, 1998. Citado 4 vezes nas páginas 42, 45, 72, and 142.

GOLEMAN, Daniel P. *Emotional intelligence: Why it can matter more than IQ for character, health and lifelong achievement*. [S.l.]: New York: Bantam Books, 1995. Citado 3 vezes nas páginas 45, 72, and 81.

GRAFT, Kris. *NetEase Revenues Rise On Strong Online Game Performance*. 2009. Disponível em: <http://www.gamasutra.com/view/news/114647/NetEase_Revenues_>. Citado na página 92.

GRIFFITHS, Mark D; DAVIES, Mark NO; CHAPPELL, Darren. Online computer gaming: a comparison of adolescent and adult gamers. *Journal of adolescence*, Elsevier, v. 27, n. 1, p. 87–96, 2004. Citado 2 vezes nas páginas 45 and 62.

GUARDASCIONE, Pietro. *The rising need for game economy designers in freemium mobile games*. 2018. Disponível em: <http://www.gamasutra.com/view/news/328645/>. Citado 3 vezes nas páginas 24, 27, and 80.

GUITART, Anna; RÍO, Ana Fernández del; PERIÁÑEZ, África. Understanding player engagement and in-game purchasing behavior with ensemble learning. *arXiv preprint arXiv:1907.03947*, 2019. Citado 2 vezes nas páginas 26 and 27.

GUITART, Anna; TAN, Shi Hui; RÍO, Ana Fernández del; CHEN, Pei Pei; PERIÁÑEZ, África. From non-paying to premium: Predicting user conversion in video games with ensemble learning. *arXiv preprint arXiv:1906.10320*, 2019. Citado na página 27.

HAGGIS-BURRIDGE, M. Four categories for meaningful discussion of immersion in video games. *Published on Researchgate*, 2020. Citado 2 vezes nas páginas 45 and 62.

HARBORD, Charly; DEMPSTER, Euan. Avatars: The other side of proteus's mirror. In: SPRINGER. *Joint International Conference on Entertainment Computing and Serious Games*. [S.l.], 2019. p. 412–416. Citado 3 vezes nas páginas 45, 61, and 128.

HARRISON, Brent; WARE, Stephen G; FENDT, Matthew W; ROBERTS, David L. A survey and analysis of techniques for player behavior prediction in massively multiplayer online role-playing games. *IEEE Transactions on Emerging Topics in Computing*, IEEE, v. 3, n. 2, p. 260–274, 2015. Citado 3 vezes nas páginas 24, 26, and 80.

HASSOUNEH, Diana; BRENGMAN, Malaika. A motivation-based typology of social virtual world users. *Computers in Human Behavior*, Elsevier, v. 33, p. 330–338, 2014. Citado na página 27.

HINGSTON, Philip. The 2k botprize. In: IEEE. *2009 IEEE Symposium on Computational Intelligence and Games*. [S.l.], 2009. p. 1–1. Citado na página 152.

HINGSTON, Philip. A turing test for computer game bots. *IEEE Transactions on Computational Intelligence and AI in Games*, IEEE, v. 1, n. 3, p. 169–186, 2009. Citado na página 152.

HOLMGÅRD, Christoffer; LIAPIS, Antonios; TOGELIUS, Julian; YANNAKAKIS, Georgios N. Evolving personas for player decision modeling. In: IEEE. *2014 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2014. p. 1–8. Citado na página 80.

HSU, Chin-Lung; LU, Hsi-Peng. Why do people play on-line games? an extended tam with social influences and flow experience. *Information & management*, Elsevier, v. 41, n. 7, p. 853–868, 2004. Citado 3 vezes nas páginas 45, 62, and 128.

HUIZINGA, Johan. *Homo Ludens. São Paulo: Ed.* [S.l.]: Perspectiva, 2000. Citado na página 24.

HUIZINGA, Johan. *Homo Ludens Ils 86*. [S.l.]: Routledge, 2014. Citado 5 vezes nas páginas 45, 72, 142, 174, and 294.

HUNICKE, Robin; LEBLANC, Marc; ZUBEK, Robert. Mda: A formal approach to game design and game research. In: *Proceedings of the AAAI Workshop on Challenges in Game AI*. [S.l.: s.n.], 2004. v. 4, n. 1, p. 1722. Citado 3 vezes nas páginas 32, 45, and 61.

HUNT, JMCV. Motivation inherent in information processing and action. 1963. Citado 3 vezes nas páginas 24, 45, and 72.

IGAMI, Mitsuru. Artificial intelligence as structural estimation: Deep blue, bonanza, and alphago. *The Econometrics Journal*, 2020. Citado na página 152.

IIDA, H. A logistic model of game's refinement. *Technical Report, Department of Computer Science, Shizuoka University*, 2003. Citado na página 93.

IIDA, Hiroyuki; TAKAHARA, Kazutoshi; NAGASHIMA, Jun; KAJIHARA, Yoichiro; HASHIMOTO, Tsuyoshi. An application of game-refinement theory to mah jong. In: SPRINGER. *International Conference on Entertainment Computing*. [S.l.], 2004. p. 333–338. Citado 2 vezes nas páginas 93 and 94.

IIDA, Hiroyuki; TAKESHITA, Nobuo; YOSHIMURA, Jin. A metric for entertainment of boardgames: its implication for evolution of chess variants. In: *Entertainment Computing*. [S.l.]: Springer, 2003. p. 65–72. Citado 4 vezes nas páginas 13, 34, 92, and 94.

JANG, Kihoon; KIM, Junwhan; YU, Byunggu. Vector-based churn prediction using neural networks in mobile games. In: IEEE. *2019 IEEE International Conference on Big Data (Big Data)*. [S.l.], 2019. p. 6081–6083. Citado na página 26.

JANSZ, Jeroen; AVIS, Corinne; VOSMEER, Mirjam. Playing the sims2: An exploration of gender differences in players' motivations and patterns of play. *New Media & Society*, Sage Publications Sage UK: London, England, v. 12, n. 2, p. 235–251, 2010. Citado 3 vezes nas páginas 45, 62, and 128.

JEON, JiHoon; YOON, DuMim; YANG, SeongIl; KIM, KyungJoong. Extracting gamers' cognitive psychological features and improving performance of churn prediction from mobile games. In: IEEE. *2017 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2017. p. 150–153. Citado 2 vezes nas páginas 28 and 74.

JERČIĆ, Petar. Arousal measurement reflected in the pupil diameter for a decision-making performance in serious games. In: SPRINGER. *Joint International Conference on Entertainment Computing and Serious Games*. [S.l.], 2019. p. 287–298. Citado 2 vezes nas páginas 45 and 61.

JOHANSSON, Anja; DELL'ACQUA, Pierangelo. Comparing behavior trees and emotional behavior networks for npcs. In: IEEE. *2012 17th International Conference on Computer Games (CGAMES)*. [S.l.], 2012. p. 253–260. Citado na página 80.

JOHANSSON, Anja; DELL'ACQUA, Pierangelo. Emotional behavior trees. In: IEEE. *2012 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2012. p. 355–362. Citado na página 80.

KELLAR, Melanie; WATTERS, Carolyn; DUFFY, Jack. Motivational factors in game play in two user groups. In: *Proceedings of DiGRA*. [S.l.: s.n.], 2005. v. 2005. Citado 3 vezes nas páginas 45, 61, and 80.

KERR, Chris. *Keywords moves into game analytics with $1.5M Yokozuna purchase*. 2018. Disponível em: <http://www.gamasutra.com/view/news/322677/Keywords_moves_into_game_analytics_with_15M_Yokozuna_purchase.php>. Citado 2 vezes nas páginas 25 and 27.

KERR, Chris. *Kixeye raises $20M to blend machine learning and game marketing*. 2018. Disponível em: <https://www.gamasutra.com/view/news/315434/Kixeye_raises_20M_to_blend_machine_learning_and_game_marketing.php>. Citado 2 vezes nas páginas 25 and 27.

KERSJES, Hanneke; SPRONCK, Pieter. Modeling believable game characters. In: IEEE. *2016 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2016. p. 1–8. Citado 2 vezes nas páginas 80 and 152.

KIM, Seungwook; CHOI, Daeyoung; LEE, Eunjung; RHEE, Wonjong. Churn prediction of mobile and online casual games using play log data. *PloS one*, Public Library of Science, v. 12, n. 7, p. e0180735, 2017. Citado 2 vezes nas páginas 25 and 89.

KIM, Yongjae; ROSS, Stephen D. An exploration of motives in sport video gaming. *International Journal of Sports Marketing & Sponsorship*, v. 8, n. 1, 2006. Citado 3 vezes nas páginas 45, 62, and 127.

KLEINGINNA, Paul R; KLEINGINNA, Anne M. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and emotion*, Springer, v. 5, n. 4, p. 345–379, 1981. Citado na página 41.

KOO, Dong-Mo. The moderating role of locus of control on the links between experiential motives and intention to play online games. *Computers in Human Behavior*, Elsevier, v. 25, n. 2, p. 466–474, 2009. Citado 2 vezes nas páginas 45 and 62.

KOSTER, Raph. *A theory of fun. 8*. [S.l.]: Aufl. Scottsdale Arizona: Paraglyph Press, 2005. Citado 2 vezes nas páginas 45 and 61.

KOZLOV, Artemiy. *Data-supported videogame shrink: what we do to understand our players better*. 2018. Disponível em: <https://www.gamasutra.com/blogs/ArtemiyKozlov/20180622/320343/Datasupported_videogame_shrink_what_we_do_to_understand_our_players_better.php>. Citado na página 150.

KRISTENSEN, Jeppe Theiss; BURELLI, Paolo. Combining sequential and aggregated data for churn prediction in casual freemium games. In: IEEE. *2019 IEEE Conference on Games (CoG)*. [S.l.], 2019. p. 1–8. Citado na página 26.

KUMAR, VISWANATHAN; SHAH, Denish. Building and sustaining profitable customer loyalty for the 21st century. *Journal of retailing*, Elsevier, v. 80, n. 4, p. 317–329, 2004. Citado na página 89.

KUMMER, Luiz Bernardo Martins; BINDER, Fábio; JR, Sylvio Barbon; NIEVOLA, Júlio Cesar; PARAISO, Emerson Cabrera. Predição do estágio de nicho em jogos rpg massivos de multijogadores utilizando o comprometimento. *Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, p. 289–300, 2016. Citado 5 vezes nas páginas 89, 91, 107, 164, and 231.

KUMMER, Luiz Bernardo Martins; NAKASHIMA, Tomoharu; THAWONMAS, Ruck; SEOK, Joon-Hong. Game analytics: Theory-based models x data-based models. *Shonan Meeting Report No. 130*, 2019. Citado 2 vezes nas páginas 28 and 33.

KUMMER, Luiz Bernardo Martins; NIEVOLA, Julio Cesar; PARAISO, Emerson Cabrera. Digital game usage lifecycle: a systematic literature review. In: *Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. [S.l.: s.n.], 2017. p. 1163–1172. Citado 4 vezes nas páginas 24, 43, 83, and 89.

KUMMER, Luiz Bernardo Martins; NIEVOLA, Julio Cesar; PARAISO, Emerson Cabrera. A key risk indicator for the game usage lifecycle. In: AAAI. *AAAI Publications, The Thirtieth International Flairs Conference (FLAIRS)*. [S.l.], 2017. p. 394–399. Citado 13 vezes nas páginas 10, 80, 88, 89, 90, 91, 92, 99, 107, 108, 164, 231, and 260.

KUMMER, Luiz Bernardo Martins; NIEVOLA, Júlio César; PARAISO, Emerson Cabrera. Applying commitment to churn and remaining players lifetime prediction. In: IEEE. *2018 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2018. p. 213–220. Citado 4 vezes nas páginas 26, 89, 98, and 267.

KUMMER, Luiz Bernardo Martins; NIEVOLA, Júlio César; PARAISO, Emerson Cabrera. Identifying niche stage in mmorpgs using ensemble classifier. In: AAAI. *AAAI Publications, The Thirty first International Flairs Conference (FLAIRS)*. [S.l.], 2018. p. 288–293. Citado 8 vezes nas páginas 26, 80, 89, 91, 107, 108, 164, and 231.

LANKVELD, Giel van. Quantifying individual player differences. *Promotores: HJ van den Herik, AR Arntz. Co-promotor: P. Spronck. Tilburg*, v. 27, 2013. Citado 6 vezes nas páginas 28, 42, 75, 162, 171, and 288.

LAWS, Robin D. *Robin's Laws of good game mastering*. [S.l.]: Steve Jackson Games, 2002. Citado 3 vezes nas páginas 45, 62, and 82.

LAZZARO, Nicole. Why we play games: Four keys to more emotion without story. 2004. Citado 2 vezes nas páginas 45 and 61.

LEE, Eunjo; JANG, Yoonjae; YOON, Du-Mim; JEON, JiHoon; YANG, Seong-il; LEE, Sang-Kwang; KIM, Dae-Wook; CHEN, Pei Pei; GUITART, Anna; BERTENS, Paul et al. Game data mining competition on churn prediction and survival analysis using commercial game log data. *IEEE Transactions on Games*, IEEE, v. 11, n. 3, p. 215–226, 2018. Citado 10 vezes nas páginas 13, 25, 26, 27, 34, 84, 89, 98, 207, and 267.

LEE, Yeng-Ting; CHEN, Kuan-Ta; CHENG, Yun-Maw; LEI, Chin-Laung. World of warcraft avatar history dataset. In: ACM. *Proceedings of the second annual ACM conference on Multimedia systems*. [S.l.], 2011. p. 123–128. Citado 3 vezes nas páginas 13, 27, and 84.

LENHART, Amanda; MADDEN, Mary; SMITH, Aaron; MACGILL, Alexandra. *Teens and Social Media*. 2007. Disponível em: <http://www.pewinternet.org/2007/12/19/teens-and-social-media/>. Citado na página 27.

LIN, Yu-Ling; LIN, Hong-Wen. A study on the goal value for massively multiplayer online role-playing games players. *Computers in Human Behavior*, Elsevier, v. 27, n. 6, p. 2153–2160, 2011. Citado 2 vezes nas páginas 45 and 62.

LOIZOU, Michael; HARTLEY, Thomas; SLATER, Sarah; NEWMAN, Robert; PANNESE, Lucia. Emotions for intelligent agents in crisis management simulations: A survey. In: IEEE. *Computer Games (CGAMES), 2012 17th International Conference on*. [S.l.], 2012. p. 213–219. Citado 2 vezes nas páginas 80 and 155.

MAKANTASIS, Konstantinos; LIAPIS, Antonios; YANNAKAKIS, Georgios N. From pixels to affect: A study on games and player experience. In: IEEE. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. [S.l.], 2019. p. 1–7. Citado 2 vezes nas páginas 28 and 74.

MALONE, Thomas W; LEPPER, Mark R. Making learning fun: A taxonomic model of intrinsic motivations for learning. *Conative and affective process analysis*, Erlbaum, 1987. Citado 2 vezes nas páginas 45 and 61.

MANSELL, Phil. *Video: Retaining player retention in Runescape*. 2015. Disponível em: <https://www.gamasutra.com/view/news/315964/Video_Retaining_player_retention_in_Runescape.php>. Citado na página 25.

MARCZEWSKI, Andrzej. Even ninja monkeys like to play. *CreateSpace Indep. Publish Platform, Charleston, Chapter User Types*, p. 69–84, 2015. Citado 3 vezes nas páginas 45, 62, and 128.

MAREN, Jean-Marie Van der. *Méthodes de recherche pour l'éducation*. [S.l.]: Presses de l'Université de Montréal et de Boeck, 1996. Citado na página 37.

MASLOW, Abraham. *Towards a Psychology of Being*. [S.l.]: New York: Van Nostrand Reinhold, 1968. Citado 5 vezes nas páginas 42, 45, 72, 142, and 205.

MASSUMI, Brian. Notes on the translation and acknowledgements. *A thousand plateaus: Capitalism and schizophrenia*, University of Minnesota Press Minneapolis, p. xvi–xixxx, 1987. Citado na página 41.

MATTHIS, Iréne. Sketch for a metapsychology of affect. *The International journal of psycho-analysis*, Routledge for the Institute of Psycho-Analysis., v. 81, n. 2, p. 215, 2000. Citado na página 41.

MCALOON, Alissa. *Analyst: 'Whale' spenders aren't as crucial to F2P mobile games as they once were*. 2018. Disponível em: <http://www.gamasutra.com/view/news/326173/Analyst_Whale_spenders_arent_as_crucial_to_F2P_mobile_games_as_they_once_were.php>. Citado na página 207.

MCALOON, Alissa. *Microsoft Studios head: 'It's really important to think about the longevity of a game'*. 2018. Disponível em: <https://www.gamasutra.com/view/news/319310/Microsoft_Studios_head_Its_really_important_to_think_about_the_longevity_of_a_game.php>. Citado na página 24.

MCALOON, Alissa. *Niantic acquires machine learning company Matrix Mill*. 2018. Disponível em: <https://www.gamasutra.com/view/news/320998/Niantic_acquires_machine_learning_company_Matrix_Mill.php>. Citado 2 vezes nas páginas 25 and 27.

MCALOON, Alissa. *World of Warcraft devs explore the history of the long-running MMO*. 2018. Disponível em: <http://www.gamasutra.com/view/news/325243/World_of_Warcraft_devs_explore_the_history_of_the_longrunning_MMO.php>. Citado na página 88.

MEHRABIAN, Albert. Basic dimensions for a general psychological theory implications for personality, social, environmental, and developmental studies. 1980. Citado 4 vezes nas páginas 45, 72, 74, and 142.

MEHRABIAN, Albert. Framework for a comprehensive description and measurement of emotional states. *Genetic, social, and general psychology monographs*, Heldref Publications, 1995. Citado 4 vezes nas páginas 45, 72, 74, and 142.

MEHRABIAN, Albert. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, Springer, v. 14, n. 4, p. 261–292, 1996. Citado 4 vezes nas páginas 45, 72, 74, and 142.

MELO, Sidney Araujo; KOHWALTER, Troy C; CLUA, Esteban; PAES, Aline; MURTA, Leonardo. Player behavior profiling through provenance graphs and representation learning. In: *International Conference on the Foundations of Digital Games*. [S.l.: s.n.], 2020. p. 1–11. Citado na página 27.

MELO, Sidney Araujo; PAES, Aline; CLUA, Esteban Walter Gonzalez; KOHWALTER, Troy Costa; MURTA, Leonardo Gresta Paulino. Detecting long-range cause-effect relationships in game provenance graphs with graph-based representation learning. *Entertainment Computing*, Elsevier, v. 32, p. 100318, 2019. Citado na página 27.

MIDDAUGH, Donna J. Managing the 80/20 rule. *MedSurg Nursing*, Jannetti Publications, Inc., v. 24, n. 2, p. 127–129, 2015. Citado na página 101.

MOORE, Geoffrey A. Crossing the chasm, 1991. *HarperBusiness, New York*, 1995. Citado 4 vezes nas páginas 10, 30, 31, and 88.

MOREIRA, Átila Valgueiro Malta; CABRAL, Danilo Rafael de Lima; CANALLE, Gabrielle Karine; ADEODATO, Paulo Jorge Leitão. A data mining solution for detection of potential buyers on mmorpgs. In: *Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. [S.l.: s.n.], 2017. p. 614–617. Citado na página 27.

MOREIRA, Márcio Borges; MEDEIROS, Carlos Augusto de. *Princípios básicos de análise do comportamento*. [S.l.]: Artmed, 2018. Citado na página 41.

MUNEZERO, Myriam D; MONTERO, Calkin Suero; SUTINEN, Erkki; PAJUNEN, John. Are they different? affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE transactions on affective computing*, IEEE, v. 5, n. 2, p. 101–111, 2014. Citado 4 vezes nas páginas 41, 42, 170, and 205.

MÜNTNER, Cindy Sheela. *The peculiar problems of the gaming industry: customer retention in MMOPRGs*. Tese (Doutorado) — Reykjavik University, 2017. Citado na página 89.

MURRAY, Henry Alexander. Explorations in personality. Oxford Univ. Press, 1938. Citado 8 vezes nas páginas 42, 45, 72, 76, 77, 79, 166, and 205.

MYERS, Isabel Briggs; MCCAULLEY, Mary H; QUENK, Naomi L; HAMMER, Allen L. *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator*. [S.l.]: Consulting Psychologists Press Palo Alto, CA, 1998. v. 3. Citado 3 vezes nas páginas 45, 72, and 142.

NACKE, Lennart E; BATEMAN, Chris; MANDRYK, Regan L. Brainhex: A neurobiological gamer typology survey. *Entertainment computing*, Elsevier, v. 5, n. 1, p. 55–62, 2014. Citado 3 vezes nas páginas 45, 61, and 128.

NAKAMURA, Jeanne; CSIKSZENTMIHALYI, Mihaly. Flow theory and research. *Handbook of positive psychology*, p. 195–206, 2009. Citado 3 vezes nas páginas 24, 45, and 72.

NAYING, Gao; PRIMANITA, Anggina; KHALID, Mohd Nor Akmal; HIROYUKI, IIDA. A key factor to maintain engagement: Case study using 'login system'. In: ATLANTIS PRESS. *Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)*. [S.l.], 2020. p. 492–497. Citado na página 26.

NEWTON, Isaac. *Philosophiae naturalis principia mathematica*. [S.l.]: London, 1687. Citado na página 92.

NORTHROP, Jackson Douglas. *Personality research form manual*. [S.l.]: research psychologists press, 1974. Citado 3 vezes nas páginas 45, 72, and 142.

NORTHROP, Jackson Douglas. *Personality research form manual*. [S.l.]: Port Huron, MI: Research Psychologists Press, 1984. Citado 3 vezes nas páginas 45, 72, and 142.

O'BRIEN, Heather L; TOMS, Elaine G. What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology*, Wiley Online Library, v. 59, n. 6, p. 938–955, 2008. Citado 3 vezes nas páginas 45, 62, and 74.

ODIERNA, Bruno Almeida; SILVEIRA, Ismar Frango. Mmorpg player classification using game data mining and k-means. In: SPRINGER. *Future of Information and Communication Conference*. [S.l.], 2019. p. 560–579. Citado na página 28.

OLSON, Cheryl K. Children's motivations for video game play in the context of normal development. *Review of general Psychology*, SAGE Publications Sage CA: Los Angeles, CA, v. 14, n. 2, p. 180–187, 2010. Citado 2 vezes nas páginas 45 and 62.

ORTONY, Andrew; CLORE, Gerald L; COLLINS, Allan. *The cognitive structure of emotions*. [S.l.]: Cambridge university press, 1990. Citado 5 vezes nas páginas 45, 72, 75, 81, and 166.

OSHRY, Barry. *Seeing systems: Unlocking the mysteries of organizational life*. [S.l.]: Berrett-Koehler Publishers, 2007. Citado na página 33.

PANUMATE, Chetprayoon; IIDA, Hiroyuki. Developing pokemon ai for finding comfortable settings. Digital Games Research Association JAPAN, 2016. Citado na página 94.

PANUMATE, Chetprayoon; IIDA, Hiroyuki. Quantifying enjoyment of individual match in games. In: *The Annual Conference on Engineering and Applied Science, Higher Education Forum*. [S.l.: s.n.], 2016. Citado na página 95.

PANUMATE, Chetprayoon; MIYAKE, Youichiro; IIDA, Hiroyuki. A generic model for emotional ai in real-time multiplayer fighting games. In: SPRINGER. *International Conference on Games and Learning Alliance*. [S.l.], 2016. p. 395–404. Citado na página 80.

PANUMATE, Chetprayoon; XIONG, Shuo; IIDA, Hiroyuki. An approach to quantifying pokemon's entertainment impact with focus on battle. In: IEEE. *Applied Computing and Information Technology/2nd International Conference on Computational Science and Intelligence (ACIT-CSI), 2015 3rd International Conference on*. [S.l.], 2015. p. 60–66. Citado na página 94.

PANUMATE, Chetprayoon; XIONG, Shuo; IIDA, Hiroyuki; KONDO, Toshiaki. Evolutionary changes of pokemon game: A case study with focus on catching pokemon. In: SPRINGER. *International Conference on Entertainment Computing*. [S.l.], 2015. p. 182–194. Citado na página 94.

PARK, Namsu; KEE, Kerk F; VALENZUELA, Sebastián. Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *CyberPsychology & Behavior*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 12, n. 6, p. 729–733, 2009. Citado na página 27.

PEDERSEN, Chris; TOGELIUS, Julian; YANNAKAKIS, Georgios N. Modeling player experience in super mario bros. In: IEEE. *2009 IEEE Symposium on Computational Intelligence and Games*. [S.l.], 2009. p. 132–139. Citado na página 28.

PERIÁÑEZ, África; SAAS, Alain; GUITART, Anna; MAGNE, Colin. Churn prediction in mobile social games: towards a complete assessment using survival ensembles. In: IEEE. *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. [S.l.], 2016. p. 564–573. Citado 3 vezes nas páginas 25, 26, and 89.

PINE, B Joseph; GILMORE, James H. *The experience economy: work is theatre & every business a stage*. [S.l.]: Harvard Business Press, 1999. Citado 2 vezes nas páginas 45 and 61.

PLUTCHIK, R. *A general psychoevolutionary theory of emotion In R. Plutchik & H. Kellerman (Eds.) Emotion: Theory, research, and experience (Vol. 1, pp. 189-217)*. [S.l.]: New York: Academic" Press, 1980. Citado 4 vezes nas páginas 45, 72, 81, and 205.

POPESCU, Alexandru; BROEKENS, Joost; SOMEREN, Maarten Van. Gamygdala: An emotion engine for games. *IEEE Transactions on Affective Computing*, IEEE, v. 5, n. 1, p. 32–44, 2014. Citado 7 vezes nas páginas 13, 75, 76, 78, 80, 160, and 166.

PRZYBYLSKI, Andrew K; WEINSTEIN, Netta; RYAN, Richard M; RIGBY, C Scott. Having to versus wanting to play: Background and consequences of harmonious versus obsessive engagement in video games. *CyberPsychology & Behavior*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 12, n. 5, p. 485–492, 2009. Citado 4 vezes nas páginas 45, 62, 126, and 127.

QUINLAN, J Ross. *C4. 5: programs for machine learning*. [S.l.]: Elsevier, 2014. Citado na página 99.

REILLY, Andrew; ROOY, Dirk Van; ANGUS, Simon. An exploration of the relationship between personality and strategy formation using market farmer: Using a bespoke computer game in behavioural research. In: SPRINGER. *Joint International Conference on Entertainment Computing and Serious Games*. [S.l.], 2019. p. 311–323. Citado 2 vezes nas páginas 45 and 62.

RÍO, Ana Fernández del; GUITART, Anna; PERIÁÑEZ, África. A time series approach to player churn and conversion in videogames. *arXiv preprint arXiv:2003.10287*, 2020. Citado 2 vezes nas páginas 26 and 27.

ROLLINGS, Andrew; ADAMS, Ernest. *Andrew Rollings and Ernest Adams on game design*. [S.l.]: New Riders, 2003. Citado 2 vezes nas páginas 86 and 87.

ROSENTHAL, Casey; CONGDON, Clare Bates. Personality profiles for generating believable bot behaviors. In: IEEE. *Computational Intelligence and Games (CIG), 2012 IEEE Conference on*. [S.l.], 2012. p. 124–131. Citado na página 80.

ROTHMEIER, Karsten; PFLANZL, Nicolas; HÜLLMANN, Joschka; PREUSS, Mike. Prediction of player churn and disengagement based on user activity data of a freemium online strategy game. *IEEE Transactions on Games*, IEEE, 2020. Citado na página 26.

RUNGE, Julian; GAO, Peng; GARCIN, Florent; FALTINGS, Boi. Churn prediction for high-value players in casual social games. In: IEEE. *2014 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2014. p. 1–8. Citado 2 vezes nas páginas 25 and 89.

RUSSELL, James A; BARRETT, Lisa Feldman. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of personality and social psychology*, American Psychological Association, v. 76, n. 5, p. 805, 1999. Citado 2 vezes nas páginas 42 and 170.

RUSSELL, James A; MEHRABIAN, Albert. Evidence for a three-factor theory of emotions. *Journal of research in Personality*, Elsevier, v. 11, n. 3, p. 273–294, 1977. Citado 4 vezes nas páginas 45, 72, 74, and 142.

SCHERER, Klaus R. Toward a dynamic theory of emotion. *Geneva studies in Emotion*, v. 1, p. 1–96, 1987. Citado na página 205.

SCHERER, Klaus R. Appraisal considered as a process of multilevel sequential checking. *Appraisal processes in emotion: Theory, methods, research*, New York, NY, v. 92, n. 120, p. 57, 2001. Citado na página 205.

SCHERER, Klaus R; SCHORR, Angela; JOHNSTONE, Tom. *Appraisal processes in emotion: Theory, methods, research*. [S.l.]: Oxford University Press, 2001. Citado na página 205.

SCHRAMM, Wilbur; LYLE, D; EDWIN, B. Television in the lives of our children. *Stan-ford: Stanford University Press. SchrammTelevision in the Lives of Our Children1961*, 1961. Citado na página 205.

SCHRÖDER, Marc. Speech and emotion research: An overview of research frameworks and a dimensional approach to emotional speech synthesis. *Doctoral thesis, Phonus 7, Research Report of the Institute of Phonetics, Saarland University*, 2004. Citado na página 155.

SHAKER, Noor; ASTERIADIS, Stylianos; YANNAKAKIS, Georgios N; KARPOUZIS, Kostas. A game-based corpus for analysing the interplay between game context and player experience. In: SPRINGER. *International Conference on Affective Computing and Intelligent Interaction*. [S.l.], 2011. p. 547–556. Citado na página 28.

SHEFFIELD, B.; ALEXANDER, L. *Focus On Korea: T3's Kim Talks Hitmaking With Audition*. 2008. Disponível em: <https://www.gamasutra.com/view/news/107795/>. Citado na página 92.

SHERRY, John L; LUCAS, Kristen; GREENBERG, Bradley S; LACHLAN, Ken. Video game uses and gratifications as predictors of use and game preference. *Playing video games: Motives, responses, and consequences*, v. 24, n. 1, p. 213–224, 2006. Citado 3 vezes nas páginas 45, 62, and 128.

SHESTOV, Anatoly. *The Game Designer's Toolbox: User stories, or how to help a player get a grasp of new content*. 2018. Disponível em: <http://www.gamasutra.com/blogs/AnatolyShestov/20181212/332084/The_Game_Designers_Toolbox_User_stories_or_how_to_help_a_player_get_a_grasp_of_new_content.php>. Citado na página 174.

SIFA, Rafet; DRACHEN, Anders; BAUCKHAGE, Christian; THURAU, Christian; CANOSSA, Alessandro. Behavior evolution in tomb raider underworld. In: IEEE. *2013 IEEE Conference on Computational Intelligence in Games (CIG)*. [S.l.], 2013. p. 1–8. Citado na página 27.

SKANSI, Sandro. *Introduction to Deep Learning: from logical calculus to artificial intelligence*. [S.l.]: Springer, 2018. Citado 3 vezes nas páginas 97, 100, and 104.

SKINNER, Burrhus Frederic. *Ciência e comportamento humano*. [S.l.]: Martins Fontes São Paulo, 2003. v. 10. Citado na página 41.

SMITH, Brian Sutton. *The ambiguity of play*. [S.l.]: Harvard University Press, Cambridge, Massachusetts, 1997. Citado 2 vezes nas páginas 45 and 61.

SMITH, Craig A; ELLSWORTH, Phoebe C. Patterns of cognitive appraisal in emotion. *Journal of personality and social psychology*, American Psychological Association, v. 48, n. 4, p. 813, 1985. Citado 2 vezes nas páginas 45 and 72.

SMUTS, Jan Christiaan. *Holism and evolution: By JC Smuts*. [S.l.]: Macmillan and Company, 1927. Citado na página 33.

SNODGRASS, Sam; MOHADDESI, Omid; HARTEVELD, Casper. Towards a generalized player model through the peas framework. In: *Proceedings of the 14th International Conference on the Foundations of Digital Games*. [S.l.: s.n.], 2019. p. 1–7. Citado na página 75.

SOWA, John F. *Conceptual structures: information processing in mind and machine*. [S.l.]: Addison-Wesley Pub., Reading, MA, 1983. Citado na página 109.

SPELLER_III, Thomas Hughes. *The business and dynamics of free-to-play social-casual game apps*. Tese (Doutorado) — Massachusetts Institute of Technology, 2012. Citado 9 vezes nas páginas 10, 24, 25, 26, 27, 88, 89, 90, and 92.

SUNDSTRÖM, Petra. *Exploring the affective loop*. Tese (Doutorado) — Stockholm University, 2005. Citado na página 28.

SUTIONO, Arie Pratama; PURWARIANTI, Ayu; IIDA, Hiroyuki. A mathematical model of game refinement. In: SPRINGER. *International Conference on Intelligent Technologies for Interactive Entertainment*. [S.l.], 2014. p. 148–151. Citado na página 94.

SWEETSER, Penelope; JOHNSON, Daniel; SWEETSER, Jane; WILES, Janet. Creating engaging artificial characters for games. In: CARNEGIE MELLON UNIVERSITY. *Proceedings of the second international conference on Entertainment computing*. [S.l.], 2003. p. 1–8. Citado 2 vezes nas páginas 80 and 205.

TAMASSIA, Marco; RAFFE, William; SIFA, Rafet; DRACHEN, Anders; ZAMBETTA, Fabio; HITCHENS, Michael. Predicting player churn in destiny: A hidden markov models approach to predicting player departure in a major online game. In: IEEE. *2016 IEEE Conference on Computational Intelligence and Games (CIG)*. [S.l.], 2016. p. 1–8. Citado 2 vezes nas páginas 25 and 89.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. Association analysis: basic concepts and algorithms. *Introduction to Data mining*, Addison Wesley Boston, MA:, p. 327–414, 2005. Citado 7 vezes nas páginas 96, 97, 98, 99, 100, 103, and 106.

THUE, David; BULITKO, Vadim; SPETCH, Marcia; WASYLISHEN, Eric. Interactive storytelling: A player modelling approach. In: *AIIDE*. [S.l.: s.n.], 2007. p. 43–48. Citado 3 vezes nas páginas 32, 45, and 61.

TIMOTHY, Ryan. *The Battle of the Gut*. 2020. Disponível em: <https://gamasutra.com/blogs/TimothyRyan/20200407/360759/The_Battle_of_the_Gut.php>. Citado na página 24.

TINBERGEN, Niko. The study of instinct. Clarendon Press/Oxford University Press, 1951. Citado na página 40.

TINBERGEN, Niko. On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, Wiley Online Library, v. 20, n. 4, p. 410–433, 1963. Citado 3 vezes nas páginas 10, 40, and 41.

TOPRAC, Paul; ABDEL-MEGUID, Ahmed. Causing fear, suspense, and anxiety using sound design in computer games. In: *Game sound technology and player interaction: Concepts and developments*. [S.l.]: IGI Global, 2011. p. 176–191. Citado 4 vezes nas páginas 45, 72, 79, and 81.

TURING, Alan M. Computing machinery and intelligence. In: *Parsing the turing test*. [S.l.]: Springer, 2009. p. 23–65. Citado na página 152.

URIARTE, Alberto. *StarCraft Brood War Data Mining*. 2017. Disponível em: <http://nova.wolfwork.com/dataMining.html>. Citado 2 vezes nas páginas 13 and 84.

VANDENBERGHE, Jason. *Developing Accurate Empathy*. 2018. Disponível em: <https://www.gamasutra.com/view/news/322074/Video_How_devs_can_apply_the_5_domains_of_play_to_develop_empathy.php>. Citado 4 vezes nas páginas 32, 45, 61, and 81.

WALDE, Sabine Schulte Im. Experiments on the automatic induction of german semantic verb classes. *Computational Linguistics*, MIT Press, v. 32, n. 2, p. 159–194, 2006. Citado na página 106.

WANG, Zhao; SAPIENZA, Anna; CULOTTA, Aron; FERRARA, Emilio. Personality and behavior in role-based online games. In: IEEE. *2019 IEEE Conference on Games (CoG)*. [S.l.], 2019. p. 1–8. Citado na página 27.

WAWRO, Alex. *Blizzard experiments with machine learning to fight Overwatch toxicity*. 2018. Disponível em: <https://www.gamasutra.com/view/news/316060/Blizzard_experiments_with_machine_learning_to_fight_Overwatch_toxicity.php>. Citado 2 vezes nas páginas 25 and 27.

WEBER, Ben. *A History of Game Analytics Platforms*. 2018. Disponível em: <https://www.gamasutra.com/blogs/BenWeber/20180409/316273/A_History_of_Game_Analytics_Platforms.php>. Citado 2 vezes nas páginas 26 and 44.

WEILLER, Thais. *Game Start: Lições de Game Design para seu videogame*. [S.l.]: Game Start, 2015. Citado 3 vezes nas páginas 45, 61, and 205.

WITTEN, Ian H; FRANK, Eibe; HALL, Mark A; PAL, Christopher J. *Data Mining: Practical machine learning tools and techniques*. [S.l.]: Morgan Kaufmann, 2016. Citado na página 99.

WU, Jen-Her; WANG, Shu-Ching; TSAI, Ho-Huang. Falling in love with online games: The uses and gratifications perspective. *Computers in Human Behavior*, Elsevier, v. 26, n. 6, p. 1862–1871, 2010. Citado 2 vezes nas páginas 45 and 62.

XIONG, Shuo; IIDA, Hiroyuki. Attractiveness of real time strategy games. In: IEEE. *Systems and Informatics (ICSAI), 2014 2nd International Conference on*. [S.l.], 2014. p. 271–276. Citado na página 94.

XIONG, Shuo; ZAHI, He; ZUO, Long; WU, Mingyang; IIDA, Hiroyuki. Analysis of the" heroes of the storm". *Advances in Computer Science: an International Journal*, v. 4, n. 6, p. 79–82, 2015. Citado na página 94.

XIONG, Shuo; ZUO, Long; CHIEWVANICHAKORN, Rachaya; IIDA, Hiroyuki. Quantifying engagement of various games. In: INFORMATION PROCESSING SOCIETY OF JAPAN. *The 19th Game Programming Workshop 2014*. [S.l.], 2014. Citado na página 94.

XIONG, Shuo; ZUO, Long; IIDA, Hiroyuki. Quantifying engagement of electronic sports game. *Advances in Social and Behavioral Sciences*, v. 5, p. 37–42, 2014. Citado na página 94.

YANNAKAKIS, Georgios N; COWIE, Roddy; BUSSO, Carlos. The ordinal nature of emotions. In: IEEE. *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. [S.l.], 2017. p. 248–255. Citado 3 vezes nas páginas 28, 30, and 48.

YANNAKAKIS, Georgios N; TOGELIUS, Julian. Experience-driven procedural content generation. *IEEE Transactions on Affective Computing*, IEEE, v. 2, n. 3, p. 147–161, 2011. Citado na página 28.

YEE, Nick. The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and virtual environments*, MIT Press, v. 15, n. 3, p. 309–329, 2006. Citado 4 vezes nas páginas 27, 45, 61, and 74.

YEE, Nick; BAILENSON, Jeremy. The proteus effect: The effect of transformed self-representation on behavior. *Human communication research*, Oxford University Press Oxford, UK, v. 33, n. 3, p. 271–290, 2007. Citado 2 vezes nas páginas 45 and 61.

YICONG, Wu; AUNG, Htun Pa Pa; KHALID, Mohd Nor Akmal; IIDA, Hiroyuki. Evolution of games towards the discovery of noble uncertainty. In: IEEE. *2019 International Conference on Advanced Information Technologies (ICAIT)*. [S.l.], 2019. p. 72–77. Citado 3 vezes nas páginas 94, 95, and 263.

ZHAO, Shiwei; WU, Runze; TAO, Jianrong; QU, Manhu; LI, Hao; FAN, Changjie. Multi-source data multi-task learning for profiling players in online games. In: IEEE. *2020 IEEE Conference on Games (CoG)*. [S.l.], 2020. p. 104–111. Citado na página 75.

ZHU, Ling; LI, Yifan; ZHAO, Guanshi. Exploring the online-game life cycle stages. In: IEEE. *E-Business and E-Government (ICEE), 2010 International Conference on*. [S.l.], 2010. p. 2436–2438. Citado 7 vezes nas páginas 24, 32, 45, 61, 88, 153, and 288.

ZILLMANN, Dolf. Mood management through communication choices. *American Behavioral Scientist*, Sage Publications, v. 31, n. 3, p. 327–340, 1988. Citado 3 vezes nas páginas 45, 72, and 142.

ZILLMANN, Dolf. Mechanisms of emotional involvement with drama. *Poetics*, Elsevier, v. 23, n. 1-2, p. 33–51, 1995. Citado 3 vezes nas páginas 45, 72, and 142.

ZILLMANN, Dolf. The psychology of suspense in dramatic exposition. *Suspense: Conceptualizations, theoretical analyses, and empirical explorations*, p. 199–231, 1996. Citado 3 vezes nas páginas 45, 72, and 142.

ZILLMANN, Dolf. Mood management: using entertainment to full advantage. In: *Communication, Social Cognition, and Affect (PLE: Emotion)*. [S.l.]: Psychology Press, 2015. p. 163–188. Citado 3 vezes nas páginas 45, 72, and 142.

ZOELLER, Georg. *Development Telemetry in Video Games Projects*. 2010. Disponível em: <http://www.gdcvault.com/play/1012227/Development-Telemetry-in-Video-Games>. Citado 2 vezes nas páginas 26 and 44.

Appendix

APPENDIX A – Supplementary Materials

All detailed information regarding the psychological models and the Concept Lattices of this work is provided in this file: Appendix_A/UEFs applications and Concept Lattices.xlsx

The Concept Lattices were built using the Concept Explorer[1]. For those with interest in navigating into the generated lattices, the input files used in this work can be found at Appendix_A/ConceptLattice_Files/

[1] For more details, please see: http://conexp.sourceforge.net/index.html

APPENDIX B – 3PIS Manual

The 3PIS manual can be found at: Appendix_B/Manual 3pis v1.pdf