Lappeenranta University of Technology
School of Business and Management
Degree Program in Computer Science

Timo Hynninen

IDENTIFYING PLAYER PROFILES IN VIDEO GAMES

Examiners: D. Sc (Tech) Jussi Kasurinen
M. Sc (Tech) Janne Parkkila

TIIVISTELMÄ

Lappeenrannan teknillinen yliopisto
School of Business and Management
Tietotekniikan koulutusohjelma

Timo Hynninen

Pelaajaprofiilien tunnistaminen videopeleissä

Diplomityö

2016

49 sivua, 7 kuvaa, 5 taulukkoa

Työn tarkastajat: TKT Jussi Kasurinen
DI Janne Parkkila

Hakusanat: pelaajatyypit, pelaajamotivaatio, pelaajaprofiilit, pelit, datatiede, hierarkinen klusterointi
Keywords: player types, player motivations, player profiles, video games, data science, hierarchical clustering

Tämä diplomityö tarkastelee pelaajatyyppien ja pelaajamotivaatioiden tunnistamista videopeleissä. Aiempi tutkimus tuntee monia pelaajatyyppien malleja, mutta niitä ei ole liiemmin sovellettu käytäntöön peleissä. Tässä työssä suoritetaan systemaattinen kirjallisuuskartoitus erilaisista pelaajatyyppien malleista, jonka pohjalta esitetään useita pelaajien luokittelutapoja. Lisäksi toteutetaan tapaustutkimus, jossa kirjallisuuden pohjalta valitaan pelaajien luokittelumalli ja testataan mallia käytännössä tunnistamalla pelaajatyyppejä data-analytiikan avulla reaaliaikaisessa strategiapelissä.

ABSTRACT

Lappeenranta University of Technology
School of Business and Management
Degree Program in Computer Science

Timo Hynninen

Identifying player profiles in video games

Master's Thesis

49 pages, 7 figures, 5 tables

Examiners: D. Sc (Tech) Jussi Kasurinen
M. Sc (Tech) Janne Parkkila

Keywords: player types, player motivations, player profiles, video games, data science, hierarchical clustering

This thesis investigates ways to identify player types or motivations in video games. Previous research identifies many approaches to player typologies but few of these approaches have been applied to games outside of limited research settings. A literature review on player typologies is conducted and several player type models are presented. Additionally a case study where one of the player typologies discovered in the literature is tested with an online real-time strategy game. As a result a classification of volunteer players and their play-motivations is built using a machine learning method called hierarchical clustering.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisors, Jussi Kasurinen and Janne Parkkila, for examining my work, even if it meant answering my calls and messages during summer holidays. The support has been invaluable.

Thanks to all of my family who have been so supportive. I’m grateful and proud to be a son, brother, uncle and a god father. A special thanks to my mother who decided to buy me a graduation present already in January even though this thesis and my degree finished in August.

“What I hear, I forget; What I see, I remember; What I do, I understand.”
– Confucius, 551-497 BC.

At Lappeenranta on July 30th 2016
- Timo T. Hynninen

TABLE OF CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

c	Cophenetic correlation coefficient
DGD	Demographic Game Design (model)
GUR	Game User Research
MMO	Massively Multi-player Online (game)
MMORPG	Massively Multi-player Online Role-Playing Game
MUD	Multi User Dungeon (game)
NPC	Non-Player Character
OCEAN	Openness, Consciousness, Extraversion, Agreeableness, Neuroticism
PENS	Player Experience of Need Satisfaction
RITE	Rapid Iterative Testing and Evaluation
RPG	Role-Playing Game
RTS	Real-time Strategy (game)
SMP	Systematic Mapping Process

1 INTRODUCTION

“Wouldn't it be great if we knew why people played social games?” [1]

Understanding how and why people play games has kept ludologists at work from well before computers and digital games right up to today. Play styles or different player personas may or may not have evolved during the past decade but the prospects of studying players have thoroughly been revolutionised by the possibilities of modern computer-era. Game developers too have for some time been interested in asking the same questions.

Over the recent years developments in game business have called for identifying different types of players and play-styles. As new business models have emerged the game business faces new challenges in terms of marketing. When games (and other digital media) are designed specifically for attracting all potential customers the understanding of different types of players and play is crucial. [2]

Now the question becomes not what makes games fun but rather how do we measure players and how do we measure what drives the players in games. No one, not the big publishers nor indie game developers want to make bad games. For game developers understanding the player is invaluable.

Game development, especially games by big publishers, is immensely data-driven [3]. The constant evolution of developing better and better data gathering tools goes hand in hand with the constant need to understand players better and better – ultimately to create better games for players. To achieve this, we must understand which particular features are liked by which particular type of players [1].

In analysing what kind of different motivations players have when playing games we might ask some of the following questions: How do players play a game? What do they get right on the first try? What do they do wrong? What kind of mistakes or sequence of actions cause the player frustration and when do players drop the game entirely? We might not only be asking how many players trigger certain events but rather why they act as they do.

In addition to gaining a deeper understanding into the motivations of players in a game we may also track the players actions while they play. This could lead to smart player tracking systems performing simple yet potentially effective actions that customise the game individually for each player and improve retention. Maybe harmful play-patterns could be detected during a game session – Being able to detect when a player is on a slippery slope could allow the game to pre-emptively guide them back to the right track.

In this thesis we study the different player profiles in video games. A literature mapping study is conducted to form a clear overview of different player typologies that have been distinguished in previous research. Additionally, an experimental case study using an online, multi-player real-time strategy (RTS) game is conducted. In the case study data from the online RTS game is collected and fitted into one of the player profiling frameworks that the literature search yielded. The applicability of the selected player profiling method is then analysed by using data from the RTS game.

1.1 Goals and delimitations

The main aspirations of this thesis are two-fold: First an overview of game analytics (and game analysis) along with a literature mapping study on player categorisation methods or frameworks is presented. The literature mapping should provide an up-to-date view of player categorisations and player typologies that have previously been identified in other studies.

Secondly, a case study using a multi-player strategy game will be employed. A player categorisation framework or a player typology will be selected based on the literature reviewed in the previous sections. Based on the player categorisation scheme selected game play metrics will be designed and utilised to collect data from the game.

The video game will be used to gather data from multi-player game sessions with a selected amount of volunteers participating. The game play data will be analysed using statistical analysis and data mining methods. Ideally the data analysis should confirm that the selected player categorisation framework or method can be successfully employed to the video game case and hence give an indication of the method's maturity and applicability to video games in general.

The main research question in this thesis is the following:

- How do we classify players to a player-category without prior knowledge about the players?

The main research question is divided into the following sub-questions:

1. What player categorization frameworks / typologies are there?
2. Are player categorizations / typologies universal or do genres differ?
3. Can we show that previously distinguished player profile models / frameworks / methods can be utilised in an online multi-player RTS game?

The goals of the thesis are following. The thesis aims to bring together player types or player categorisations identified in previous research and numerical analysis of game-play data. In summary the main contributions of this thesis are the following:

- Background and understanding of game analytics, game user research, metrics and their importance for business intelligence is presented.
- A literature mapping study of player types and how player types are distinguished from games is conducted.
- A classification of players in a real time strategy game using a player typology model is done using data analytics tools.

The thesis recognises the following delimitations:

- The player categorisation scheme is only tested on one game.
- Psychological factors behind player motivations are not in the scope of the work.

- A new classification or typology model will not be constructed. Instead a player type model from previous research will be taken as a starting point and it's suitability in classifying players in video games will be tested in a case study.

1.2 Structure of the thesis

The rest of this thesis is structured as follows.

In section 2 background on game analysis is presented and terms of the game user research (GUR) field are clarified. Related work in the field of player typologies and player categorisation by playing motivations is discussed.

In section 3 a systematic literature mapping study is presented. The section focuses on the mapping study, it's search protocol and results. In the end, an overview of all articles found in the search is presented.

In section 4 a case study where game player profiles are identified from an open source multi-player game using a player type model selected from the results of the literature review. Game-play data is subjected to machine learning analysis and groups of players are distinguished. Additionally, the resulting player profiles are examined.

Finally, in section 5 discussion, conclusions and summary of the contributions in this thesis are presented.

2 METHODOLOGY OVERVIEW

This section goes over the research process, different research methods that are employed and topics used in the latter part of the thesis. First the research process that is employed is defined and presented. Game analysis as a field of science along with the qualitative and quantitative measures employed in the field are discussed in the following two subsections. The fourth subsection covers the area of clustering in general as a tool for dissecting data collected from video games.

2.1 Research process

The whole research process consists of six phases and it is visualised in figure 1. In phase 1 the research problem is formulated and formalised. While the research problem is clarified research questions are formed to complement, unclutter and form delimitations.

In phase 2 the research process itself is planned. All the steps in the research process are locked in and preliminary plans for each step are formed. This process should result in the formalisation of later steps in the research process.

In phase 3 the research area is clarified. The result of this step should form a clear overview into the different research and application areas in which this work is situated. The purpose of this phase is to form the related work and research methodology sections of this thesis.

In phase 4 a literature mapping study is carried out. The mapping study follows the formal process of carrying out systematic literature reviews in software engineering by Petersen et al. [4].

In phase 5 the resulting literature is analysed and synthesised. An overview of the included articles and their topics is presented. Additionally a time line of the articles the search yielded is presented.

Finally, in phase 6 a player typing scheme is selected from the literature and applied small scale in a case study. The case study uses an open source online multi-player game played by a group of people, while game-play metrics are being recorded. The metrics will then be normalised, clustered and analysed by hand.

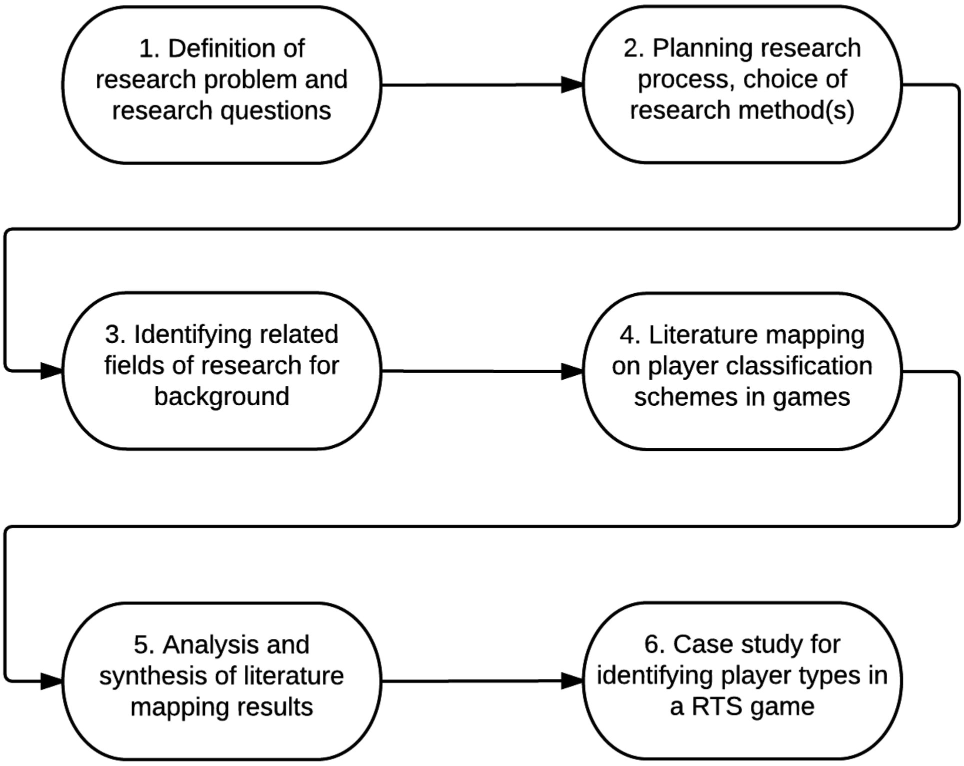

Figure 1: Research process outline.

2.2 Introduction to game analytics

Game analysis is an emerging business and an interesting study subject: Game development is becoming more and more data-driven nowadays. That is to say games are being developed in a player centric way, constantly analysing player behaviour in-game. Developers use analysis of player data to make informed decisions across all levels of the game development process. Also in the highly competitive game business it is vital for developers and publishers to know their users as well as possible. Analysis tools are used to gain better knowledge of players. [3]

Game analytics has been a game-changer for game developers over the past few years. A widespread adoption of data-driven business intelligence practices can be seen as a cornerstone of the game industry's working methods. Currently, quantitative user-oriented game research methods are paramount. The free-to-play business model for one has been a force driving the need to understand players better and better. [3]

Game user research (GUR) methods, like any other research methods, divide into qualitative and quantitative categories. Most of the common GUR methods are qualitative. For example, methods called think-aloud or RITE (Rapid Iterative Testing and Evaluation) are heavily user-centric. In RITE a single comment from a single test subject may incur immediate changes to the game. [5]

However, also quantitative methods such as play-testing in large numbers or A/B testing can be employed in GUR. These methods rely on quantitative data collected from the game using hooks and signals inside the game's program code. These data are then analysed using statistics. Needless to say that quantitative data is of much narrower domain compared to qualitative data but it can reveal the most common flaws in game design where the game design and user expectation fail to meet. [5]

Like almost any field of research, both quantitative and qualitative research methods can be employed and combined. Qualitative game user research is about developing different player metrics and collecting data. The focus is on the data analysis of different metrics collected from one player (or a group of single players). Collecting quantitative data from players in small numbers is relatively easy. It can be done for example by interviews, biometrics or by simply observing players whilst they are playing.

Quantitative player research on the other hand, focuses on a multitude of players providing data from one (or a limited number) of metrics. These metrics are typically collected from the players' in-game actions, such as game-play states, user-interface interaction or in-game decisions. Telemetry systems for game developers make the collection of quantitative data from all their customers easy.

2.3 Tracking player behaviours in games

In recent years, many game companies – from indie game developers to big publishers – have started to collect game telemetry. Telemetry is data obtained over a distance. This can, for example, be quantitative data about how a user plays a game, tracked from the game client and transmitted to a collection server. [3], [6]

A key feature of games is that underneath they are essentially state machines. This applies to both digital games (video games) and non-digital games (board games, card games etc.) [3]. During a play session the player performs actions and responds to prompted inputs. These actions form a loop wherein the game state changes [3], [7]. Games can thus be described as deterministic finite automata.

A state machine describing a game has a finite amount of states regardless of whether it has a set end state. The fact that games underneath the surface are state machines means that it is easy for a game analyst or a game user researcher to track player progression and performance in a game. This is done by examining the current state of the game at each point in time.

Another way of analysing player behaviour based on the different game states is to focus not on the game states themselves but transitions between game states. Each transition is then an event that contains information of the players progression. In addition to the progression metric we can also track the path a player has taken through the game by considering what the previous and following states are.

Tracking player behaviour in games allows the game developers to collect data either locally in play-testing research or over distance using telemetry systems. Additionally, player tracking may also be used to construct feedback loops in games. Feedback loops, also sometimes referred to as cybernetic systems, are systems that monitor themselves and the surrounding systems and if necessary, make changes [7]. This can be particularly useful in games for modifying game-play conditions and for example, adjust the difficulty setting for the player [8].

2.4 Metrics in game-play analysis

"Game metrics are interpretable measures of something related to games. More specifically, they are quantitative measures of attributes of objects. A common source of game metrics is telemetry data of player behaviour." [3]

Metrics are quantifiable measures taken from a players game session. They are collected by actively monitoring a player's game session and actions in a game. That is to say, metrics are collected either by having a test player in a monitored facility or by incorporating telemetry tools inside a game.

Game analysis tools allow developers to gather play data from a number of players remotely. A single metric is a measurement that indicates what happens in a player's game session to the game developer or a GUR researcher. Game-play metrics are related to the behaviour of a game user (player) inside a game. These metrics are collected to gain insight on game design and user experience. Game-play metrics can either be frequency based or event based.

Frequency based metrics are measures collected from the game at a given interval. Frequency based recording of telemetry is generally used when the attribute of the object being tracked is always present. For example, a player character in a virtual world always has a position. The player's movements can then be examined by examining the players location at certain intervals.

Event based metrics on the other hand are used to collect information on state changes inside the game. These metrics can include any interaction the player has with the game, including but not limited to user interface events (for example, starting a game session), interacting with other users in an online game (socialising or chat), movement within the game world, or anything related to the mechanics of the game (such as reaching an objective or performing a task).

2.5 Numerical analysis of game-play metrics

Gathering data using different metrics can be laborious but fortunately it can be automated using telemetry systems. Analyzing what the data indicates however requires some method of data analytics. In this thesis we employ an unsupervised machine learning method called clustering.

Clustering is the task of grouping a set of objects in such a way that objects in the same group, a cluster, are more similar (in one sense or another) to each other than to those in other clusters. It is a basic method of exploratory data mining, and a common technique for statistical data analysis. Clustering methods are used in many fields: For example machine learning, pattern recognition, image analysis, information retrieval or bioinformatics. [9]

As illustrated in figure 2, two or more objects belong to the same cluster if they are close to each other according to a given distance. In this example the type of clustering is distance based, where the distance metric is simply the distance between two points in space. The goal of clustering is to determine the internal grouping in a set of previously unlabelled data. [10]

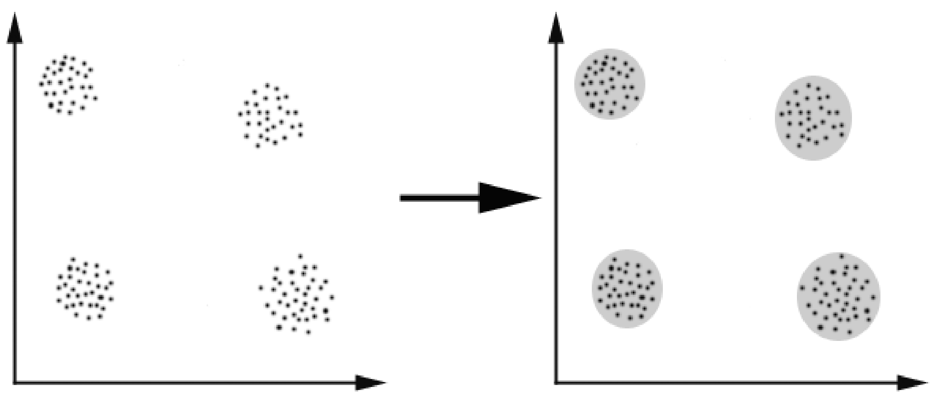

Figure 2. Clusters (highlighted in grey) formed from raw data. [10]

Cluster analysis does not refer only to one specific method or algorithm. Instead the term is used to denote the general problem of grouping samples in a dataset without any training data that would usually be a requirement for building a classifier. Clustering can be achieved by various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently they find them. A cluster means simply a group of data objects. However, different cluster models can be used for different ends and means and clustering algorithms vary between clustering models.

Typical cluster models include connectivity-based clustering, centroid-based clustering, distribution-based clustering and density-based clustering. Later on in this thesis, in the following chapters, connectivity-based clustering will be used as the clustering method for game data analysis. Following is a summary of different clustering methods and a reasoning for choosing a connectivity-based method later on.

Connectivity-based clustering (hierarchical clustering) is based on the idea of connecting objects to form clusters based on their distance to other objects. A cluster is described by the maximum distance needed to connect parts of the cluster. Hierarchical clustering does not take into account clear outlier data points. That is to say there is no notion of noise involved which is a drawback of using the method. [11]

In hierarchical clustering there is no single division of the data set. Instead, clustering algorithms provide a hierarchy of clusters that merge with each other at certain distances. The result of a connectivity-based clustering algorithm is a hierarchy of subsets of data, from which we must select the appropriate groups. [12]

This hierarchical cluster tree is represented using a plot called a dendrogram [13]. A sample dendrogram is depicted in figure 3. In a dendrogram, the y-axis marks the distance at which the clusters merge, while the objects are placed along the x-axis such that the clusters don't mix. Depending on how much pruning of the cluster tree we want to apply the number of resulting clusters varies.

In the dendrogram in Figure 3 starting from the stem of the tree we can vary between 1 to 4 levels of pruning to result from 2 to 10 clusters of objects. If we set pruning to stop the clustering at the first link level we would have two clusters of 7 and 3 objects in each. If not pruned at all (level 4) the end result will place each of the 10 objects into their own cluster. When analysing the cluster tree the user must decide the level of pruning, depending on what kind of clusters emerge at which link level.

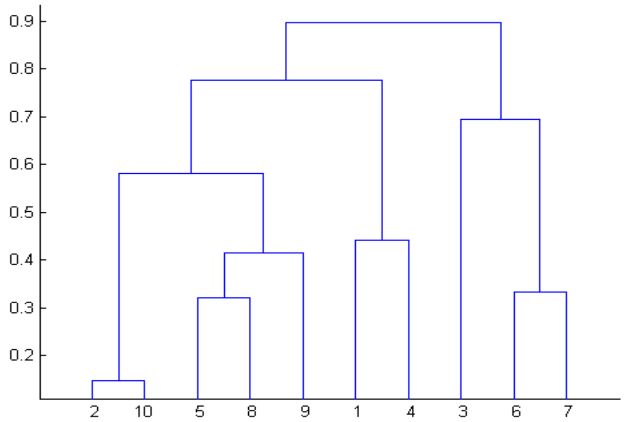

Figure 3. A dendrogram [13].

Connectivity-based clustering is a family of methods that differ by the way distances between objects are computed [9]. A common metric for calculating the distance between two objects is the Euclidean distance:

$$|a-b|=\sqrt{\sum_i \left(a_i-b_i\right)^2}$$

where *a* and *b* are points in an n-dimensional space and p1 .. pn are the Cartesian coordinates of the points [p =(p1, p2, ... pn)]. Euclidean distance is the ordinary distance between two points, i.e. a straight line.

Other popular methods of clustering data are centroid-based clustering and distribution-based clustering [9]. In centroid-based clustering, clusters are represented by a central vector, which may not necessarily be a member of the data set. The algorithm finds cluster centres and assigns each objects to the nearest cluster centre, so that the squared distances from the cluster are minimized. However, centroid-based clustering cannot represent density-based clusters.

Distribution-based clustering models are most closely related to statistics based on distribution models. Clusters can then easily be defined as objects belonging most likely to the same distribution. Distribution-based clusters suffer from one key problem which is over-fitting.

Over-fitting in statistics and machine learning means that the statistical model which is supposed to represent the data input describes noise or randomness instead of the underlying relationships between data points. In general, over-fitting is when the data model is too descriptive of a single data set to be generally descriptive. [14] This will lead to the clustering method not being generalisable.

In density-based clustering clusters are defined as areas of higher density than the remainder of the data set. Objects in these sparse areas, which separate cluster from each other, are usually considered to be noise and border points. The drawback of these clustering algorithms is that they expect a density drop to detect cluster borders.

In this thesis we will focus on the hierarchical clustering method. This is due to our particular research setting, where cluster detection does not need to be autonomic. To determine the validity (goodness) of our clustering method we will use the cophenetic correlation coefficient metric [12]. Cophenetic distance is a measure of how similar objects have to be in order to be grouped into the same cluster [15]. The cophenetic correlation coefficient *c* is calculated using the following formula

$$\frac{\sum_{i<j}\left(x(i,j)-\bar{x}\right)\left(t(i,j)-\bar{t}\right)}{\sqrt{\left[\sum_{i<j}\left(x(i,j)-\bar{x}\right)^2\right]\left[\sum_{i<j}\left(t(i,j)-\bar{t}\right)^2\right]}}$$

where

- *x(i, j)* is the Euclidian distance between two points (*i* and *j)*
- *t(i, j)* is the dendrogrammatic distance between the same points in the cluster model.

2.6 Player typologies

This subsection discusses the different approaches to divide players into groups depending on their personality types. A personality typology or a model of personality types refers to the psychological classification of different types of individuals. In player profiling, previous research has focused on identifying player types and characteristics, yet it has rarely connected these characteristics with game components [16].

In psychology there are several different models for personality types. The classification of people dates back to ancient Greece and the Hippocratic theory of four humours based on the dominance of bodily fluids (blood, phlegm, black bile and yellow bile) [17]. Modern psychology knows personality type models such as the Five-Factor personality structure [18] or Keirsey Temperaments [19]. In this thesis these personality typologies are out of scope however, as we are primarily interested in personality types in video games.

Probably the most well known player personality categorisation is the Hearts, clubs, diamonds and spades model by Bartle [20] depicted in figure 4. The common names describing the underlying player motivation in the categorisation are killers, socializers, achievers and explorers. Other player profile categorisations have since been proposed, some of which build on the Bartle model.

The player types "arise from the inter-relationship of two dimensions of playing style: action versus interaction, and world-oriented versus player-oriented." The categorization was realized after an extensive study of Multi-User Dungeon (MUD) games. A MUD is a multi-player real-time virtual world. Usually text-based, MUDs combine elements of role-playing games, hack and slash, player versus player, interactive fiction, and online chat. [20]

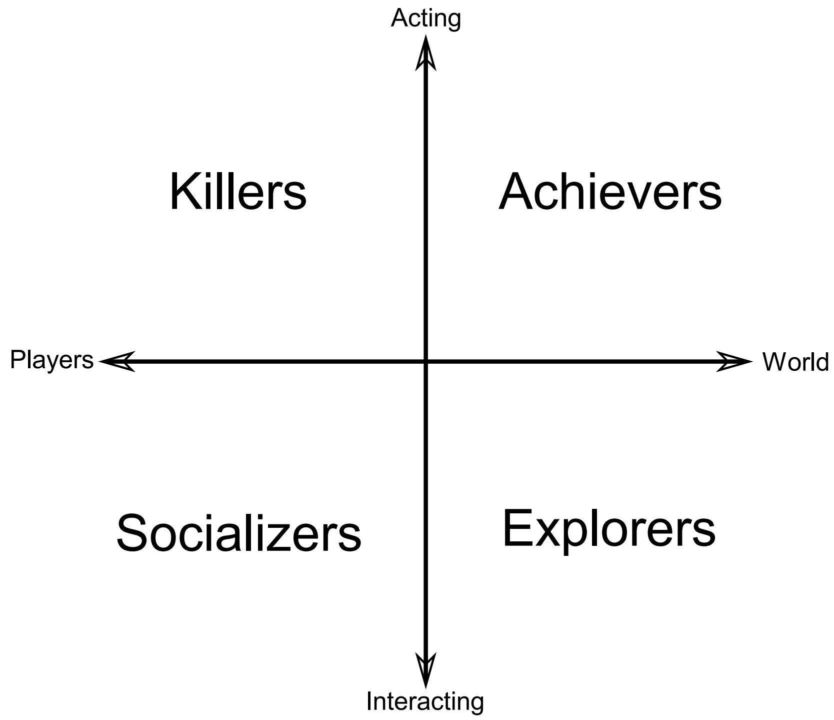

Figure 4: Bartle's player taxonomy [21].

Bartle asked questions such as are MUD games similar in nature to sporting games or board games, pastimes like reading or gardening, sports such as hunting or fishing, entertainment like nightclubs or television? The four field classification is formed as a result from analysing these questions.

According to Bartle, the first type of players are Achievers. These are players who prefer to gain points, levels, equipment and other concrete measurements of succeeding in a game. They will go to great lengths to achieve rewards that confer them little or no game-play benefit simply for the prestige of having it.

Explorers are players who prefer discovering areas, creating maps and learning about hidden places. They often feel restricted when a game expects them to move on within a certain time, as that does not allow them to look around at their own pace. They find great joy in discovering an unknown glitch or a hidden Easter egg.

There are a many players who choose to play games for the social aspect, rather than the actual game itself, called Socializers. They gain the most enjoyment from a game by interacting with other players, and on some occasions, computer-controlled characters with personality. The game is merely a tool they use to meet others in-game or outside of it.

The final player type is Killers. Killers enjoy competition with other players. Killers especially prefer fighting other players to non-player character opponents.

Bartle also created an extended, 8-type version of his player model for virtual world players. These types are: Friend, Griefer, Hacker, Networker, Opportunist, Planner, Politician and Scientist. Many studies, such as [22], base their work on Bartle's original four-field model or use it as a starting point for creating their own classification.

Bartle's player types have previously been used for example in profiling employees [23] and finding interaction profiles of software engineering students [24]. The model has been moderately criticised. Some criticism stem from the fact that player types based on thirty-year old games are too specific in scope to be applicable to anything else [25]. Bartle himself states that the model is not specific and very generalisable but does at the same time it may be hard to apply the model to a specific end [1].

Different player typologies from previous research have previously been identified in a literature meta-synthesis study by Hamari and Tuunanen in [26]. Next their results are summarized and the relevant player typologies are presented. Later on in this thesis results of another systematic literature mapping study are presented on player types is presented and a quick synthesis of it's results is introduced.

In general, the meta-synthesis by Hamari and Tuunanen concentrates on the different classification schemes of players based on psycho-graphic factors. Hamari & Tuunanen reviewed the different ways "in which players have been typified in previous relevant literature." The authors formed a meta-synthesis of the different typologies in order to "clarify the current state of research and suggest further research avenues."

The authors state that marketing practises of segmentation and differentiation are nowadays a part of game design: Segmentation is the process of identifying different customer groups and differentiation is identifying and serving the needs of a customer group [27]. The argument is that games and game design have to take into account who their potential customer is and hence understanding player types and playing motivations is an essential ability for game publishers.

Distinguishing between types of players and play styles help achieve the aforementioned goal. Hamari & Tuunanen note, how previous studies have concentrated on distinguishing characteristic psycho-graphic and behavioural aspects of game-play. Especially gamer's motivation and in-game behaviour has been covered extensively. The first approach to categorising players presented in [26] is a psycho-graphics based solution that often manifests in the division of players into hard-core and casual players. This hard-core v. casual players categorisation is swiftly debunked for being too simplistic and not necessarily even a distinction that can always be made.

The second approach to play style is behavioural basis. This kind of categorization includes the four Bartle types (killer, socializer, achiever, explores) [20] or the four play styles in *Tomb Raider: Underworld* (veteran, solver, pacifist, runner) [28]. Yee's motivational factors [29] build on the Bartle types but are not exactly player types (but rather play motivations). The authors note how that Yee's model takes into account both behavioural and psycho-graphic factors and interplay between them. Other approaches to play styles exists but they do not conform to the two umbrella terms of behavioural or psycho-graphic basis.

3 PLAYER TYPES IN GAMES

The purpose of this section is to answer the following questions about player typologies:

- What player classification frameworks and player typologies have been distinguished in previous literature?
- How applicable are the existing methods for classifying players in games.

To gain an overview about the breadth of research in this field in order to answer these questions a literature mapping study will be conducted. The rest of this chapter will go over the literature mapping method, conducting the actual search of literature and the result of the search, and a brief overview of the topics that emerge from the acquired literature. The literature review is aimed to uncover practical applications for player typologies along with background research into the different motivations players have toward games.

3.1 A systematic mapping study

In order to acquire relevant literature on player typologies a literature search process was conducted. The protocol for this search process follows the systematic mapping process by Petersen et al. [4] which describes how to perform systematic mapping studies in the domain of software engineering.

The systematic mapping process (SMP) consists of six individual phases. Different steps and their outcomes are the following:

1. **Definition of research questions**
   The first step of a SMP addresses the main goal of a literature study, which is to provide an overview of the research area and identify the research within it. Additionally literature mapping identifies the forums in which research in the given scope is published. The research questions for the study should be formed to reflect both these goals.
   Outcome: Scope of the review (research). Research questions.

2. **Conducting the search**
   In this step the search strings to be used with selected databases are formed and the searches are conducted. It is not necessarily required to use the same search strings for all databases.
   Outcome: All articles that came up using the selected search strings.

3. **Screening of papers**
   The third step of the SMP process is to form inclusion and exclusion criteria which are then used in screening all the articles found. Inclusion (exclusion) criteria determine whether the articles are relevant (or not) for the study. The research questions should be used as a starting point for determining these clusivity rules.
   Outcome: Relevant articles

4. **Key-wording using abstracts**
   Keywords are used to develop the classification scheme, by which the processed articles are then sorted.
   Outcome: Classification scheme

5. **Data extraction and mapping process**
   When the classification scheme is in place relevant articles can be sorted into the classification for further analysis. This data extraction phase requires reviewing each of the screened articles by placing them first into the keyword based classification and consequently forming new classes based on the articles as they are evaluated. This means that the classification scheme is constantly being updated based on the content of the reviewed articles. The resulting classification is then presented in either table or diagram form.
   Outcome: Systematic map

In this literature search this process is slightly modified because of the nature of the literature search: Since the goal of the search is to find different player typologies, classification methods and understand their applicability to online games instead of analysing the thematics of state-of-art academic research, part of the last step in the systematic literature mapping process [4] will not be conducted and left for future work. Also in steps 1-3 the fora in which the found articles are published will not be extensively reported for the same reasons.

3.2 Conducting the mapping study

Conducting the literature mapping study begins with a search of article literature. The search will be two fold: first scientific databases (Web of Science, Digra Digital Library, ACM Digital Library, IEEE Xplore and ScienceDirect) followed by a complementary google scholar search. The mapping study aims to answer the following questions:

- What articles will come up in the search?
- What methods for distinguishing player types have been researched?
- Have player typologies / player categorisations been previously applied and how?

Search strings that are used are variations of ”player type” and ”player motivation” with also the plural forms of the words, ”types” and ”motivations”

The search protocol is as follows: First we perform the searches in the scientific databases. The titles of all resulting articles will be screened, first by title and keywords, then by abstract and finally the entire text if the article is not excluded by title, keywords or abstract.

Inclusion criteria and exclusion criteria are used to decide, whether an article should be included or not. The inclusion criteria are the following:

- Any article where a player type model or player characterisation framework is presented.
- Any article, where a player type model or characterisation framework has been applied.
- Any article that presents different player motivations in video games.

Exclusion criteria are the following:

- Any article behind a pay-wall or one where only a citation is available.
- Any article where the scope is not video games.

Results of the searches are presented in table 1. An overview of the articles is presented in the next sub-section. All in all, a total of 18 articles were uncovered from the scientific databases.

Table 1. Results of the literature searches.

Search Place / Search Terms	***"player type(s)"***	***"player motivation"***	**Total**
Web of Science	3 / 107	3 / 14	6 / 121
Digra DL	0 / 0	2 / 6	2 / 6
ACM DL	4 / 7	2 / 4	6 / 11
IEEE Xplore	1 / 4	2 / 13	3 / 17
ScienceDirect	0 / 2	1 / 13	1 / 15
TOTAL	**8** / 120	**10** / 50	**18** / 170
Complementary searches	**8** (/ 4500 (approx.)	**5** (/ 1500 approx.)	**13** (/ 6000+)
Total after complementary searches based on the results			**31** (/ 6000+)

Complementary searches using the Google and Google Scholar search engines were also performed during the search process. However, due to the sheer size of the search space it would have required massive effort to produce even a small amount of new results. For this reason the Google searches were not entirely exhaustive – processing of the search results was stopped after tens of titles stopped having any meaningful relation to the search terms.

First off, the same search strings used with the main scientific databases were entered to Google Scholar to perform a broad, complementary search. Additionally, searches were made during the paper screening phase to find citations or related articles to already included articles. These searches yielded additional articles of which 11 were eventually included into the results.

3.3 Overview of articles acquired in the literature search

This section presents a broad overview of all the articles that were acquired in the literature searches. A brief summary of each article is presented. Additionally, a time line of the articles is shown in table 2.

[30] discusses the player typologies needed to study player satisfaction. The article goes over traditional psychometric typologies such as Bartle [20] or Yee's player motivations [29] and dissects their faults as generalisable models. The article presents results from questionnaires aimed to study player motivations and suggests that in order to form a robust, universal player typology player's playing preferences (trait theory) will have to be taken as a starting point. Building on their own previous work on player profiles, the authors present a demographic game design model (DGD) where player's are classified into four archetypes: Conqueror, Manager, Wanderer and Participant

[31] investigates churn prediction (players not returning to the game) in RPGs. The authors compare player clustering with a data driven approach with many features and a player motivation based approach with few features. The results show that the motivational features are better suited for predicting player behaviour.

[32] investigates the psychogenic needs of game players in six different categories: materialism, power, affiliation, achievement, information and sensual needs. The authors state that each of the six needs provoke certain actions which can be used as variables in game-play analysis. The study does not discuss why people play games but rather how they are motivated in the game. After creating the classification for game-play motivations and their relations to in-game actions the author matches the framework by evaluating it using the RPG game, The Witcher.

[33] applies a psychogenic needs framework to a game and investigates the relations between different needs and game mechanics behind them. The needs framework by Murray [34] deals with the psychological needs of humans, ranging from materialistic, power, affiliation, achievement, information and sensual needs. The game used in the study is an RPG, Fallout 3. The study concludes that the Murray’s framework is applicable to gaming because player motivations can be linked to different game-play situations.

[35] highlight the requirement for games to be more responsive to different player types and their individual needs. The underlying message is that games should have the capability of adapting to the player. The study suggests games should constantly monitor and model the player and adjust their setting accordingly.

[8] describes an approach to player-centered game design. It divides player-centric game design into three subgroups of related areas: understanding players, modelling players and adaptive game technologies. The study discusses the possibilities of player-centricity, such as adapting a game's challenge level, smoothing learning curves and enhancing game-play experience individually for players.

[36] presents a machine learning approach to categorising Pac-Man players. A real-time classifier using Decision tree algorithms is used resulting in good (70 %) accuracy. The players are categorised into the Demographic Game Design model (GDG) by Bateman et al. ([30] covered earlier on in this summary).

In [37] player behaviour is studied by using in-game geographic information as the metric. Players' movement is analysed using maps, reports and charts in Tomb Raider: Underworld action/adventure game. The study concludes that spatial analysis may be used in user experience research as well as a testing tool.

In [28] work on player types in Tomb Raider: Underworld is continued by distinguishing four typical player types in the game using clustering and self-organizing maps. The player types are Veterans, Solvers, Pacifists and Runners.

[38] presents a comparison of clustering methods used in player segmentation. In the study behavioural telemetry from over 70 000 World of Warcraft users is clustered using different clustering algorithms. The applicability of the different clustering methods is reviewed and the consequences of choosing a particular clustering tool are presented.

[39] uses Bartle's player types to show how players consume products in games or virtual worlds. It highlights the differences between consumerism in real-world and game-world economies. The study is limited to player motivations in virtual world economics.

[40] assesses player motivations and learning strategies based on their personality. The objective of the study is to provide guidelines for game designers of learning games.

[41] investigates personalised gamified systems. The article proposes a possible model to link player type, personality traits and types with game element and mechanic combinations.

[42] aims to quantify players' involvement in games based on Bartle's player types. The study uses a 3D exploration game as means of collecting data. According to the results players of this particular game have two roles, either explorers or achievers.

[43] presents an approach to player types based on mastery and performance in gaming. The classification is a shallow, achievement related four-field: super-achievers, mastery-only, performance-only and non-achievers.

[44] present a qualitative method for analysing video games based on game activities. The authors used three puzzle games and a Player Experience of Need Satisfaction (PENS) survey to measure player experience in the games. The method can be utilized to evaluate the types of activities in a game and act as a tool for balancing game design.

[45] presents a "cross-genre, cross-cultural, behaviourally validated scale" for motivations in game play. In the study a large scale survey (18 000 respondents) was issued to massively multi-player online (MMO) gamers. As a result, the article synthesises a taxonomy of six player profiles: Socializer, completionist, competitor, escapist, story-driven, and smarty-pants.

[46] explores the relationship between personality and game player types. The authors conducted a survey aimed at people on internet gaming forums to gather information about people's personality and play behaviours. The study establishes connections between

personality type theory and the demographic game design model. Although the study does not use the player categories to classify players of any particular game the categorization is universal between genres.

[47] investigates how player motivation, player performance, and player enjoyment are connected in EverQuest II, a multi-player RPG. The results are formulated by measuring game data and linking the game data to player motivation by survey results. The authors suggest future study should investigate in-game behaviours and identify behavioural patterns that can correspond to player motivations.

[48] applies a previously formed player personality profile model ([49]) to a real, commercial video game (Fallout 3). In the study metrics are collected from the game and they are correlated to the players' personality profiles. The study results in a successful effort to use game metrics to uncover player profiles in a video game.

[16] presents an experiment setting which applies categories of psychogenic needs to players actions when playing a platformer. The of psychogenic motivators in the categorisation coined in the study are Competition, Challenge, Enjoyment, Social Interaction, Diversion, Fantasy Interests, Arousal or Excitement and Entertainment.

[50] goes over multiple psychological systems of player categorisations to try and arrive at a unified model in which player behaviour can be understood. The unified model consists of a synthesis between Bartle's four player types, the four Keirsey Temperaments (as categorised by psychologist David Keirsey [19]), and Bateman’s DGD [30]. The classification of players in this unified model is: Artisan / Killer / Experientialist, Guardian / Achiever / Gamist, Rational / Explorer / Simulationist and Idealist / Socializer / Narrativist.

[51] identifies player types on Massively multi-player online games. The classification relies on a small number of features and consists of three player types: Killers put the highest priority on attacking monsters, Plan Agents focus on progressing in the game world. Markov Killers are something between the two, calculating what the best next move would be.

[52] collect metrics from Hitman: Blood Money game to form play-personas which model how players interact with the game. The uncovered play styles are particular to the game in question and the metrics collected from game sessions are specific rather than generic.

[49] investigates whether a personality profile can be determined by observing a player's behaviour in a game. The authors use a model of five personality traits to model a player profile. In order to validate the model an RPG game was used to gather data from test participants' game play.

The study measured how various different game-play metrics correlated to the personality traits. Authors used five personality traits from the Five Factor Model of personality (OCEAN): Openness, Consciousness, Extraversion, Agreeableness, Neuroticism. Metrics that were gathered from the game were total movement in game, conversations with non-player characters, number of finished games, movement per area and conversation per NPC (non player character). The results showed correlation between video game data and personal traits. The study concludes that a video game can be used to create an adequate personality profile of a player.

[53] concludes that motivations in Massively Multi-player Online Games consist of Sociability, Competition, Immersion, Achievement, and Challenge. These findings are in line with previous research, notably with Bartle's player types.

[54] distinguishes five different player types in relation to motivation, behaviour and influence on others in multi-player health games. The player types are: Achievers, Active Buddies, Social Experience Seekers, Team Players, and Freeloaders. The scope of the study is quite specific as the article concentrates on health games for the youth.

[29] presents an empirical model of player motivations in online games. The model consists of 10 motivational components (advancement, mechanics, competition, socializing, relationship, teamwork, discovery, role-playing, customisation and escapism) that can be grouped under three umbrella categories (achievement, social and immersion components).

Table 2. Time line of the included articles from the literature search.

Year	Article name	Author(s)
2003	Identification of player types in massively multiplayer online games	Thawonmas et al.
2004	Dynamic player modeling: A framework for player-centered digital games	Charles & Black
2005	Player-centred game design: Player modelling and adaptive digital games	Charles et al.
2006	Motivations for play in online games	Yee
2007	Virtual consumption: Using player types to explore virtual consumer behavior	Drennan & Keeffe
2008	Defining Personas in Games Using Metrics	Tychsen & Canossa
2009	Player profiling and modelling in computer and video games	Cowley
2009	Explorations in Player Motivations: Game Mechanics.	Bostan et al.
2009	Player modeling using self-organization in Tomb Raider: Underworld	Drachen et al.
2009	Analyzing spatial user behavior in computer games using geographic information systems	Drachen & Canossa
2009	Explorations in Player Motivations: Game Mechanics	Bostan et al.
2009	Player Motivations: A Psychological Perspective	Bostan
2010	Assessing Players' Motivations and Learning Strategies Based on their Personality	Felicia
2011	Personality And Play Styles: A Unified Model	Stewart
2011	The differences in motivations of online game players and offline game players: A combined analysis of three studies at higher education level	Hainey et al.
2011	Player typology in theory and practice	Bateman et al.
2011	Beyond Player Types: Gaming Achievement Goal	Heeter et al.
2011	Games as personality profiling tools	Van Lankveld et al.
2011	Churn Prediction in MMORPGs Using Player Motivation Theories and an Ensemble Approach	Borbora et al.
2011	An Exploratory Study of Player Performance, Motivation, and Enjoyment in Massively Multiplayer Online Role-Playing Games	Shim et al.
2012	Player Profiling with Fallout 3.	Spronck et al.
2012	Personality and Player Types in Fallout New Vegas	McMahon et al.

Table 2. Continued.

2012	This is not a one-horse race: understanding player types in multiplayer pervasive health games for youth	Xu et al.
2013	Towards personalised, gamified systems: an investigation into game design, personality and player typologies	Ferro et al.
2013	Real-time rule-based classification of player types in computer games	Cowley et al.
2013	Motivation During Videogame Play: Analysing Player Experience in Terms of Cognitive Action	Inchamnan & Wyeth
2014	Towards Quantifying Player's Involvement in 3D Games Based-on Player Types	Hanna et al.
2014	A comparison of methods for player clustering via behavioral telemetry	Drachen et al.
2014	Player types: A meta-synthesis	Hamari & Tuunanen
2014	Player Motivations in Massively Multiplayer Online Games	Voulgari et al.
2015	The Trojan Player Typology: A cross-genre, cross-cultural, behaviorally validated scale of video game play motivations	Kahn et al.

3.4 Player motivation frameworks

In order to classify players in a game in the next section of the thesis, a choice has to be made to pick a player type model to be used in the classification. This subsection will present the motivation behind choosing Yee's player motivation model [29]. A good starting point would be the Bartle taxonomy [20]: Even though the categorisation into killers, socializers, achievers and explorers should be simple and universal, the model does not really dictate how the classification should be done. It does not mandate what kind of player actions belong inherently to which player type.

This is the same with many other classifications, such as the five traits of personality model used by van Lankveld et al. [49] or the demographic game design model by Bateman [30]. These player typologies are coined by inspecting the motivations of players but they do not cover what kind of in-game actions would be in relation to which motivational aspect.

Yee's player motivation model distinguishes the three main categories for player motivation: The senses of achievement, social interaction, and immersion. Within these main categories 10 sub-categories are laid out. Additionally, Yee's model presents typical in-game behaviours that go hand in hand with the motivations they are linked to. The model can thus be used to analyse player actions in a game without worry that there is any bias in the interpretation of which actions are to be associated with which motivation component. Yee's motivation model, with all the components and game actions are presented in table 3.

Another distinction that Yee's motivation model has to other player models that the categorisation is two dimensional. When applying the motivation model to a game it is possible to use either the specific motivation (sub-components), the main motivation (main component), or a combination of both. This gives the researcher enormous freedom, as one does not have to confine themselves within one category. For example, it may be hard to label actions to be either in the Discovery category or in the Role-Playing category. Instead the action can simply be said to belong in the immersion component.

Table 3. Yee's Player motivation categories [29].

Achievement	**Social**	**Immersion**
Advancement Progress, Power, Accumulation, Status	**Socializing** Casual Chat, Helping Others, Making Friends	**Discovery** Exploration, Lore, Finding Hidden Things
Mechanics Numbers, Optimization, Templating, Analysis	**Relationship** Personal, Self-Disclosure, Find and Give Support	**Role-Playing** Story Line, Character History, Roles, Fantasy
Competition Challenging Others, Provocation, Domination	**Teamwork** Collaboration, Groups, Group Achievements	**Customization** Appearances, Accessories, Style, Color Schemes
		Escapism Relax, Escape from Real World, Avoid Real World Problems

4 PLAYER TYPES: CASE 0 A.D.

This section presents a case study, where Yee's player motivations [29] are used to distinguish player types in an open source, multi-player real-time strategy game called 0 A.D. Metrics are collected from the game and they will be sorted into groups by using the hierarchical clustering machine learning method. Next a brief overview of the 0 A.D. game is presented. The section continues with the data analysis (clustering) and finally, presents the results of the player categorisation using the clustering method.

4.1 0 A.D. THE GAME

0 A.D. [55] is a free, open source, multi-platform multi-player real-time strategy (RTS) game. It is a historical war and economy game focusing on the years between 500 B.C. and A.D. 500 [56]. 0 A.D. is an Age of Empires II inspired RTS game where the aim is in building a town, training an army, fighting with other players or artificial intelligence, and technology research [57]. In short, the game is about economic development and warfare in the player's small, early common era civilisation [56].

Due to their flexibility in customisation open source games are also popular in academic research. 0 A.D. in particular has been used in different fields of research and teaching although not very extensively. An example application area where this open source game has been used is the simulation of night vision methods in video (and video games) [58].

The 0 A.D. game is set literally in the year 0 (Anno Domini). The player is given some resources (food, wood, rock and metal), a small number of non-player characters (units) and possibly some buildings (units). From there on, the player has to start building a civilisation, starting from a small village and (hopefully) building an empire. In order to achieve this the player needs to maintain a balance of the three most important game mechanics: Gathering resource, building units or buildings and fighting other players.

Gathering resource is done by ordering units (non-player characters, NPC) to work, for example to cut wood, farm fields or mine rock. In order to improve gathering performance, more units can be built. However, building units will cost resources. There are also different types of units: War units are specifically used for battling against other players whereas for example citizen soldiers can do both war efforts and gather resource. Different units can be built in specialized buildings: Buildings are also units and cost resources to build.

The third mechanic has to do with the game's ultimate goal: Building a widespread empire and perishing all other civilisations. As with all civilisations, the player must keep an army to protect and serve the inhabitants of their settlement. A village of a civilisation in the game is depicted in figure 5.

Figure 5: ”Gallic Fields.” A screen-shot from the 0 A.D game. [59]

Previously in section 3 we justified why Yee’s player motivation model is applicable to be used with a game with the intention of identifying different types of players. The model describes what kind of in-game actions indicate which player motivation. In table 4 the game mechanics or actions you may perform in-game are listed along with their equivalent categories in Yee's player motivation framework.

This way we can collect metrics straight from the game that indicate different player motivations. As we are limited to in-game data, two of the categories, Customization and

Escapism need to be ignored as they will not be possible to monitor: Customizable properties do not really exist in 0 A.D. and Escapism can not be monitored at all, as we are only monitoring the player's inside the game and do not have an insight into the player's psyche.

Table 4. Yee's player motivation components and corresponding in-game actions which can be monitored.

Main components of Yee's player motivation categorisation	Metrics collected from 0 A.D during a game session
Advancement	Collecting resource (Farming, wood cutting, mining).
Mechanics	Building varied units (hero units, priests). Building dedicated units (maritime units, siege weapons). Use of extraterritorial buildings or units.
Competition	Economy and military growth.
Socializing	Chat messages.
Relationship	Alliances. Treaties. Trade routes. Resource donation to allies.
Teamwork	Alliances. Trade routes.
Discovery	Exploration. Collecting treasures. Building extraterritorial buildings or units.
Role-Playing	Inventive player names. Building special buildings that do not help the main goal of the game. Playing with many different civilisations (Only if multiple game sessions are played).
Customisation	*The game does not contain customizable features (except maybe for game developers).*
Escapism	*Cannot be measured from in-game statistics.*

4.2 CASE STUDY RESEARCH PROCESS AND METHODOLOGY

For this study we recruited 11 volunteers to play 0 A.D. in a controlled and monitored usability laboratory. In the end a total of 20 game sessions were carried out by the participants with some volunteers taking part in several sessions. Although the laboratory is monitored 24/7 game sessions were not recorded using video cameras in the room or by using screen capturing software. Instead only metrics from the 0 A.D. game were collected during these game play sessions. All participants were male, aged between 22 and 35, and all of the participants identified themselves as gamers or said video games were their hobby.

The game sessions were of varying length, spanning between 25 minutes and 4 hours of game time. It should be noted that due to the nature of the game, session lengths are seldom the same for all players: Losing early on in a game would lead to a player dropping out considerably quicker compared to the remaining players who then continued until victory conditions were met.

The participants had varying skills in RTS gaming in general. Also some participants were already familiar with 0 A.D. itself whereas others were completely new to the game. However the game was picked up by everyone quickly and even though skill levels varied the players never showed signs of extreme frustration.

Because there are more recorded game sessions than participants, the less experienced players were given training rounds. The results from these games were not used in the end. The resulting data contains the telemetry from 12 different game sessions by 11 participants. Two sessions by one of the players who was most familiar with the game were put into the data set. The goal was to have these sessions produce similar results. This would serve as an additional sanity metric, so that the in the resulting clusters the same player would not be classified too far off from himself.

The data collected from the game are simple numeric scores that indicate how well the players have performed in which are of the game. For example, a player might have a ”30 000” score in collecting some individual resource and ”90 000” in resource collection combined. Scores are being collected for each of the game’s main economics: (resource) gathering, military, exploration, building (and destroying or losing) units, as well as alliances and trade between players. The scores among each other are balanced and in thousands in the order of magnitude. The scores sum up to form a total score of approximately 150 000 each game.

4.3 CASE STUDY RESULTS

The game data consists of raw integer numbers (between 0 and approximately 70 000) depicting how well the players have succeeded in the main goals of the game (resource collecting, unit building and fighting against other players). Additionally there are some percentages that show how successful a player's war efforts have been, how much of the game area was explored by each player, and how successful players were in market trading (building economical growth). In addition to these metrics, which are provided by the game itself, during the sessions also social interactions were monitored. This would include alliances between players and in-game chat.

The scores for metrics by each player was saved and labelled with names 'P1' – 'P10' for the participants. The volunteer two of the game sessions belong to is labelled 'PC'. These data were then clustered using Orange Canvas Data Mining software. Matlab was used to calculate key figures, for example the cophenetic correlation coefficient, so that we could be certain that the clusters produced were not similar in characteristics and that the cluster tree was consistent.

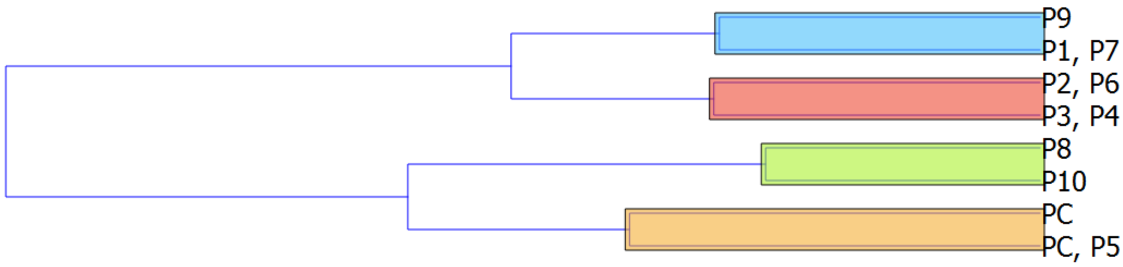

Figure 6. Player clusters with maximum depth of cluster set to 3. The width of the dendrogram is approximately 60 000 units in 0 A.D. player score.

As illustrated in figure 6, hierarchical clustering using the Euclidean distance metric produced eight distinct clusters of players with at least one player in each cluster. Pruning was used with 3 as a depth setting. Going higher in the depth setting will result to four individual clusters indicated by different highlight colours in the dendrogram. We will consider these four highlighted clusters the main clusters with the option to look deeper into the smaller subgroups within the main groupings.

After the clustering was complete, we examined the typical actions for each player cluster. We took note of the most prominent action types from the players. These were actions that the metrics indicated were used by players the most or were given a significantly better score by comparison to all other metrics. Then these actions and their archetypes were combined. As a result, each player cluster is now paired with player motivations that are characteristic to that cluster. Table 5 presents the clusters and the most principal characteristics collected from the game according to Yee's player motivation framework.

Table 5. Clusters, cluster members, and common actions and motivations the players exhibited during play.

Cluster ID	Common action types indicated by metrics	Members
C1 (P9, P1, P7)	Role-playing (Story Line, Character History) Discovery (Exploration)	3
C2 (P2, P6, P3, P4)	Socializing (Chat, Helping Others) Relationship (Find and Give support) Role-playing (Story Line, Character History)	4
C3 (P8, P10)	Advancement (Progress, Accumulation) Competition (Challenging others) Socializing (Chat, Helping Others) Relationship (Find and Give support) Teamwork (Collaboration)	2
C4 (PC, P5)	Advancement (Progress, Power, Accumulation) Mechanics (Numbers, Optimization) Competition (Challenging others, Provocation, Domination)	2

From this classification we can distinguish different player groups for all the clusters. C4 is the cluster for the most veteran players. The players gained high scores in all metrics that show progression in the game's key mechanics.

C3 consists of players who are not as familiar with the game but exhibit traits that indicate competitiveness and use of team tactics to be competitive in the game. One possible explanation for these players' actions during their game sessions may have been, that they felt teamwork could be a key strategy against more skillfull opponents.

Players in C2 scored high in metrics that indicated role-playing, socializing and in-game interaction with other players. These players were learning about the game's mechanics actively. They were keen to try out all what is possible to do in the game making their end result scattered and non-decisive.

Finally, players in C1 were the most unique in terms of metrics collected. They were the only group that actively explored the game world, often at the expense of other game mechanics. Even though all players in the study did understand the end goal of the game (which is progression, building and perishing others) these players would not care about winning too much. For them, the immersion of exploring a virtual world was their biggest interest.

In order to verify the dissimilarity of the clusters, we can inspect the resulting dendrogram (clustering result) and calculate the cophenetic correlation coefficient for the distances between different links. The dendrogram is formed so that the length of a link represents the distance (dissimilarity) between the two clusters. Also the distance between a parent and child links (the distance from one link stem to another) should be the same throughout the graph, showing that cluster dissimilarity within the graph is consistent [12].

By looking at Figure 6 we can see that the first two clusters (C1, C2) are nearly as equally far away from the link base. The other two clusters (C3, C4) are further apart from each other, indicating slightly more similarity (and not dissimilarity) between the cluster properties. Indeed, in Table 5 these two clusters (C3, C4) show the same action types. The clusters are however sufficiently apart from each other.

To verify that the clusters are indeed consisting of similar objects between dissimilar clusters, the cophenetic correlation coefficient was calculated and resulting value was 0.86. The correlation coefficient is a ratio, so the closer the value gets to 1 the better [12]. With only a small set of data 0.86 is an acceptable ratio – however changing the clustering method by for example changing the distance metric might improve on the result.

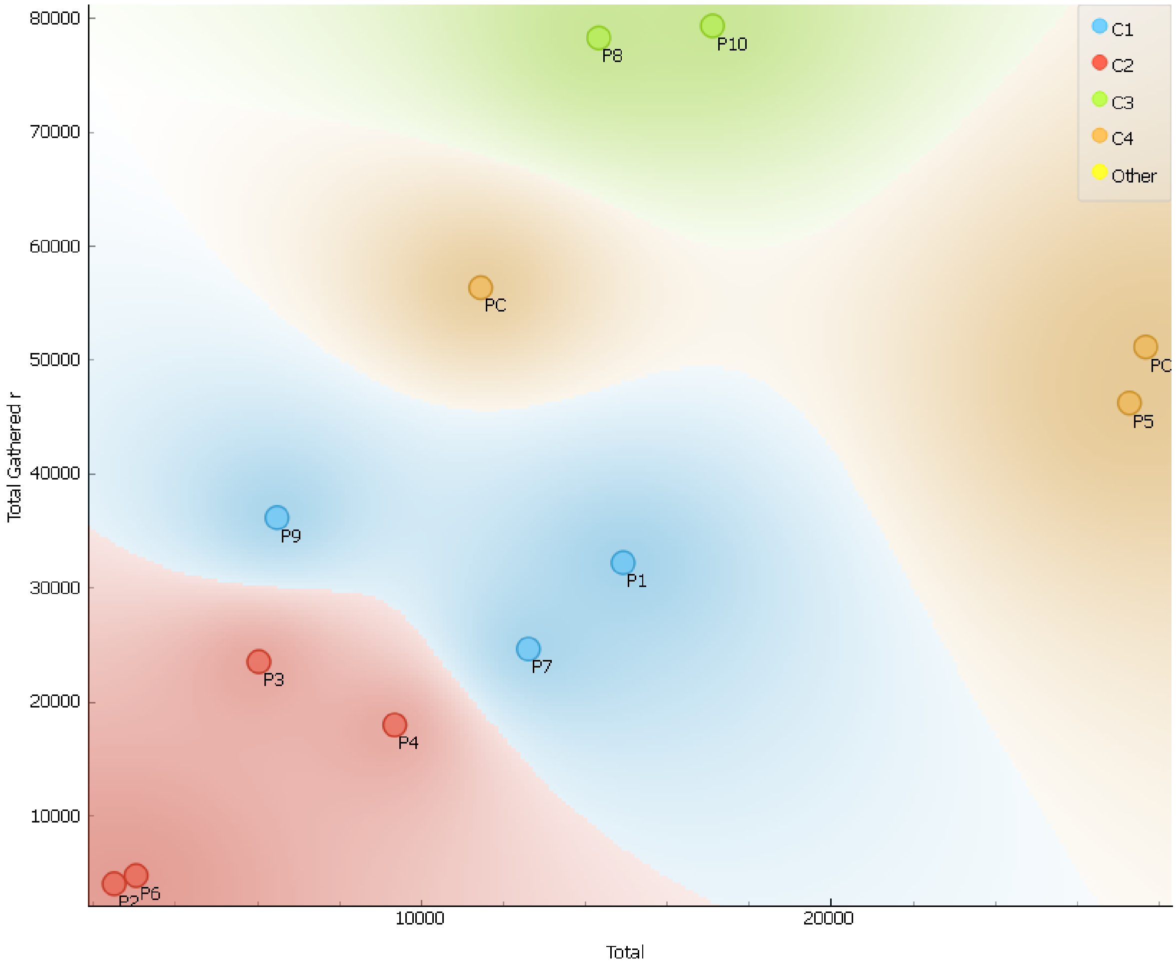

Figure 7. Classifier lines in a scatter plot with the player's total score on X-axis and score of gathered resources on Y-axis. The colour highlights are the same as for figure 6.

In figure 7 the total score is plotted in comparison to the resource gathering metric (advancement in Yee terms) to show what is the most significant (most dominant) metric of all. The gathered resources (advancement) metric is the only one which produces cubic classifier functions for these data. For all other metrics (for example building units or trade) in relation to the total score the classifiers would be clearly higher order functions.

This means that these advancement metrics in this game may be slightly more weighted. However, individual players' scores in C2, C1 and C4 show that some data points while close together are not part of the same cluster. This is an indication that this metric is not the only influential one.

5 DISCUSSION AND CONCLUSION

The collected game metrics are well separable. As we can see from examining the formed clusters and their cophenetic distances to other clusters, the clusters have dissimilar enough properties to be differentiated from each other. That is to say that by collecting data using metrics that follow Yee's player motivation framework, we can distinguish individual players or player groups from these data.

The 0 A.D. game was used in order to test whether a real multi-player RTS game would conform to the player motivation model's categorisation of motivations for in-game actions. The model seems to be robust enough to be used as a guideline for interpreting multidimensional game-play data. The model's biggest advantage over other player typology frameworks is that it's components are specific enough to be paired with various metrics collected from a game, yet the main components are general enough to fit both the genre the model was created for and the genre we tested the model on: Massively multi-player online role playing games (MMORPG) and real time strategy games (RTS).

The scoring system in 0 A.D. is obviously well tested by the open-source community around the game. It isn't inconceivable to say that much of the success in the results of the clustering is partly down to this well thought-out scoring mechanism. However, the scores are inherently collected for a different purpose than profiling players, and the player profiles have certainly not been designed with scoring in an RTS game in mind.

The scoring system is designed so that players could look at individual areas of the game and find out how different playing strategies affect the game. The scores are also accumulated to for one total score that should place the players in order from best to last. In this sense the scores are meant to separate individual players from each other, rather than finding similarities which we are effectively doing by clustering the data. Albeit that high scores in the same individual fields do indicate a similar play style the metrics have never been designed to really indicate that: These metrics are rather used to balance game play between different play strategies.

Clusters formed from game-play data are descriptive and we managed to classify players based on their in-game actions. The metrics employed worked well with the clustering. Saving game-play data using these metrics meant the data could be used to classify players by archetype and we can form conclusions from these data about the players thinking, motivations, traits, wants and needs.

Recognising player types from game-play metrics could result into predicting the players' actions or movements beforehand. It is clear that there is much potential for analysing players' in-game actions. Future work could include identifying a player's play-style from as early on in the game as possible.

The player could thus be guided towards the goals of the game if desired. For example, if the player displayed traits indicating exploring and treasure hunting, the game could generate more this kind of content to support the player's internal motivations. If exploring the game world and gathering did not play a significant role in the game's goals, or was only one of the key components in the game, the player could be subtly instructed to explore other game mechanics using messages or maybe even rewards.

Identifying the player's archetype and internal motivations could also bring the game and game developer more means of offering the right kind of stimuli in-game: For example, someone responding well to immersion and discovery traits could be offered more map to explore. An achievement hunter on the other hand could enjoy additional quests or achievements to complete.

Automating the player type identification process could be a new page-turner in game analytics. This way game user research could effectively be crowd sourced: If the tools to analyse the actions of a player in game deployed with games, data analytics with masses could be employed.

All in all, every new piece of information game developers about their audience is valuable. There is no indication that game analysis methods employed within the game industry were diminishing in any way.

6 SUMMARY

In this thesis background information and understanding of game analytics, game user research, metrics and their importance for business intelligence was presented. A literature mapping study of player types and how player types are distinguished from games was conducted. Finally, a classification of players in a real time strategy game using a player typology model was done using data analytics tools.

In previous research much effort has been put into understanding players and their motivations during game-play. Games themselves are highly measurable, both locally and over the world wide web. Even though game developers collect many metrics from games there are few examples in literature of measuring player motivation or distinguishing between player types from video games.

We presented a literature review of previous research on player types and motivations. We present different player motivation categorisations, practical approaches to utilising these models and discuss the potential use cases for distinguishing between different player types in games. From the different player typologies and categorisations we selected Yee's player motivation model as most flexible to use in a case study, where data from an online real time strategy game was used to distinguish different types of players.

The case study employs a machine learning method called hierarchical clustering in order to search for similarities between players whose play styles are alike, and dissimilarities between players who do not exhibit similar traits or actions in-game. Clustering placed 11 volunteer player's play styles in four different groups, each with their own distinct properties. For example, teamwork oriented players were distinguished from players who wanted to succeed in the game on their own.

REFERENCES

1. R. A. Bartle, “Design Principles,” in *Multiplayer: The Social Aspects of Digital Gaming*, vol. 3, Routledge, 2013, p. 10.
2. J. Hamari and J. Tuunanen, “Player types: A meta-synthesis,” *Trans. Digit. Games Res. Assoc.*, vol. 1, no. 2, 2014.
3. A. Drachen, M. Seif El-Nasr, and A. Canossa, “Game Analytics – The Basics,” in *Game Analytics*, M. Seif El-Nasr, A. Drachen, and A. Canossa, Eds. London: Springer London, 2013, pp. 13–40.
4. K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, “Systematic mapping studies in software engineering,” in *12th International Conference on Evaluation and Assessment in Software Engineering*, 2008, vol. 17.
5. H. Desurvire and M. S. El-Nasr, “Methods for Game User Research: Studying Player Behavior to Enhance Game Design,” *IEEE Comput. Graph. Appl.*, vol. 33, no. 4, pp. 82–87, Jul. 2013.
6. T. V. Fields, “Game Industry Metrics Terminology and Analytics Case Study,” in *Game Analytics*, M. Seif El-Nasr, A. Drachen, and A. Canossa, Eds. London: Springer London, 2013, pp. 53–71.
7. K. Salen and E. Zimmerman, “Chapter 18: Games as Cybernetic Systems - Feedback loops,” in *Rules of play: Game design fundamentals*, MIT press, 2004.
8. D. Charles, A. Kerr, M. McNeill, M. McAlister, M. Black, J. Kcklich, A. Moore, and K. Stringer, “Player-centred game design: Player modelling and adaptive digital games,” in *Proceedings of the Digital Games Research Conference*, 2005, vol. 285, p. 100.
9. S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press, 2008.
10. M. Matteucci, “Matteo Matteucci’s home page.” [Online]. Available: http://chrome.ws.dei.polimi.it/index.php/Matt%27s_Home_Page. [Accessed: 29-Jul-2016].
11. M. Ester, H. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise,” 1996, pp. 226–231.

12. “Hierarchical Clustering - MATLAB & Simulink - MathWorks Nordic.” [Online]. Available: http://se.mathworks.com/help/stats/hierarchical-clustering.html. [Accessed: 27-May-2015].
13. “Dendrogram plot - MATLAB dendrogram - MathWorks Nordic.” [Online]. Available: https://se.mathworks.com/help/stats/dendrogram.html?requestedDomain=se.mathworks.com. [Accessed: 27-Jul-2016].
14. B. Everitt, *Cambridge dictionary of statistics*. Cambridge University Press, 1998.
15. R. R. Sokal and F. J. Rohlf, “The comparison of dendrograms by objective methods,” *Taxon*, pp. 33–40, 1962.
16. R. Staewen, P. Trevino, and C. Yun, “Player Characteristics and their Relationship to Goals and Rewards in Video Games,” 2014.
17. I. Johansson and N. Lynøe, *Medicine & philosophy: A twenty-first century introduction*. Walter de Gruyter, 2008.
18. J. M. Digman, “Personality Structure: Emergence of the Five-Factor Model,” *Annu. Rev. Psychol.*, vol. 41, no. 1, pp. 417–440, 1990.
19. D. Keirsey and M. M. Bates, *Please understand me*. Prometheas Nemesis, 1984.
20. R. Bartle, “Hearts, clubs, diamonds, spades: Players who suit MUDs,” *J. MUD Res.*, vol. 1, no. 1, p. 19, 1996.
21. “File:Character theory chart.svg,” *Wikipedia, the free encyclopedia*. [Online]. Available: https://en.wikipedia.org/wiki/File:Character_theory_chart.svg. [Accessed: 22-Jul-2016].
22. N. Yee, “Motivations of play in MMORPGs,” 2005.
23. F. Caron, “On Mapping Bartle’s Player types To Employees,”. [Online]. Available: http://frankcaron.com. [Accessed: 5-Jan-2015].
24. A. Knutas, J. Ikonen, D. Maggiorini, L. Ripamonti, and J. Porras, “Creating Software engineering student interaction profiles for discovering gamification approaches to improve collaboration,” in *Proceedings of the 15th International Conference on Computer Systems and Technologies*, 2014, pp. 378–385.
25. Ipsy Kyatric, “Bartle’s Taxonomy of Player Types (And Why It Doesn’t Apply to Everything) - Tuts+ Game Development Article,” *Game Development Tuts+*,

2013. [Online]. Available: http://gamedevelopment.tutsplus.com/articles/bartles-taxonomy-of-player-types-and-why-it-doesnt-apply-to-everything--gamedev-4173. [Accessed: 12-May-2015].

26. J. Tuunanen and J. Hamari, "Meta-synthesis of player typologies," in *Proceedings of Nordic Digra 2012 Conference: Games in Culture and Society, Tampere, Finland*, 2012.
27. P. Kotler and K. L. Keller, *Marketing Management*, 14 edition. Harlow: Pearson Education, 2011.
28. A. Drachen, A. Canossa, and G. N. Yannakakis, "Player modeling using self-organization in Tomb Raider: Underworld," in *Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on*, 2009, pp. 1–8.
29. N. Yee, "Motivations for play in online games," *Cyberpsychol. Behav.*, vol. 9, no. 6, pp. 772–775, 2006.
30. C. Bateman, R. Lowenhaupt, and L. E. Nacke, "Player typology in theory and practice," *Proc. DiGRA Think Des. Play*, 2011.
31. Z. Borbora, J. Srivastava, K.-W. Hsu, and D. Williams, "Churn Prediction in MMORPGs Using Player Motivation Theories and an Ensemble Approach," in *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom)*, 2011, pp. 157–164.
32. B. Bostan, "Player Motivations: A Psychological Perspective," *Comput Entertain*, vol. 7, no. 2, p. 22:1–22:26, Jun. 2009.
33. B. Bostan, U. Kaplancali, K. Cad, and A. Yerlesimi, "Explorations in Player Motivations: Game Mechanics.," in *GAMEON*, 2009, pp. 5–11.
34. H. . Murray, *Explorations in personality*. Oxford, England: Oxford Univ. Press, 1938.
35. D. Charles and M. Black, "Dynamic player modeling: A framework for player-centered digital games," in *Proc. of the International Conference on Computer Games: Artificial Intelligence, Design and Education*, 2004, pp. 29–35.

36. B. Cowley, D. Charles, M. Black, and R. Hickey, “Real-time rule-based classification of player types in computer games,” *User Model. User-Adapt. Interact.*, vol. 23, no. 5, pp. 489–526, 2013.
37. A. Drachen and A. Canossa, “Analyzing spatial user behavior in computer games using geographic information systems,” in *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*, 2009, pp. 182–189.
38. A. Drachen, C. Thurau, R. Sifa, and C. Bauckhage, “A comparison of methods for player clustering via behavioral telemetry,” *ArXiv Prepr. ArXiv14073950*, 2014.
39. P. Drennan and D. A. Keeffe, “Virtual consumption: Using player types to explore virtual consumer behavior,” in *Entertainment Computing–ICEC 2007*, Springer, 2007, pp. 466–469.
40. P. Felicia, “Assessing Players’ Motivations and Learning Strategies Based on their Personality,” in *Proceedings of the 4th European Conference on Games-Based Learning: ECGBL 2009*, 2010, p. 87.
41. L. S. Ferro, S. P. Walz, and S. Greuter, “Towards personalised, gamified systems: an investigation into game design, personality and player typologies,” in *Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death*, 2013, p. 7.
42. N. Hanna, D. Richards, M. Hitchens, and M. J. Jacobson, “Towards Quantifying Player’s Involvement in 3D Games Based-on Player Types,” in *Proceedings of the 2014 Conference on Interactive Entertainment*, New York, NY, USA, 2014, p. 26:1–26:10.
43. C. Heeter, Y.-H. Lee, B. Medler, and B. Magerko, “Beyond Player Types: Gaming Achievement Goal,” in *Proceedings of the 2011 ACM SIGGRAPH Symposium on Video Games*, New York, NY, USA, 2011, pp. 43–48.
44. W. Inchamnan and P. Wyeth, “Motivation During Videogame Play: Analysing Player Experience in Terms of Cognitive Action,” in *Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death*, New York, NY, USA, 2013, p. 6:1–6:9.

45. A. S. Kahn, C. Shen, L. Lu, R. A. Ratan, S. Coary, J. Hou, J. Meng, J. Osborn, and D. Williams, “The Trojan Player Typology: A cross-genre, cross-cultural, behaviorally validated scale of video game play motivations,” *Comput. Hum. Behav.*, vol. 49, pp. 354–361, Aug. 2015.
46. N. McMahon, P. Wyeth, and D. Johnson, “Personality and Player Types in Fallout New Vegas,” in *Proceedings of the 4th International Conference on Fun and Games*, New York, NY, USA, 2012, pp. 113–116.
47. K. J. Shim, K.-W. Hsu, and J. Srivastava, “An Exploratory Study of Player Performance, Motivation, and Enjoyment in Massively Multiplayer Online Role-Playing Games,” in *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom)*, 2011, pp. 135–140.
48. P. Spronck, I. Balemans, and G. Van Lankveld, “Player Profiling with Fallout 3.,” in *AIIDE*, 2012.
49. G. van Lankveld, P. Spronck, J. Van den Herik, and A. Arntz, “Games as personality profiling tools,” in *2011 IEEE Conference on Computational Intelligence and Games (CIG)*, 2011, pp. 197–202.
50. B. Stewart, “Personality And Play Styles: A Unified Model,” *Gamasutra Www Gamasutra Com*, vol. 1, 2011.
51. R. Thawonmas, J.-Y. Ho, and Y. Matsumoto, “Identification of player types in massively multiplayer online games,” in *Proc. the 34th Annual conference of International Simulation And Gaming Association (ISAGA2003), Chiba, Japan*, 2003, pp. 893–900.
52. A. Tychsen and A. Canossa, “Defining Personas in Games Using Metrics,” in *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, New York, NY, USA, 2008, pp. 73–80.
53. I. Voulgari, V. Komis, and D. G. Sampson, “Player Motivations in Massively Multiplayer Online Games,” in *2014 IEEE 14th International Conference on Advanced Learning Technologies (ICALT)*, 2014, pp. 238–239.
54. Y. Xu, E. S. Poole, A. D. Miller, E. Eiriksdottir, D. Kestranek, R. Catrambone, and E. D. Mynatt, “This is not a one-horse race: understanding player types in

multiplayer pervasive health games for youth," in *Proceedings of the ACM 2012 conference on computer supported cooperative work*, 2012, pp. 843–852.

55. "Wildfire Games." [Online]. Available: http://wildfiregames.com/. [Accessed: 29-Jul-2016].

56. J. McElroy, "The Joystiq Indie Pitch: 0 A.D.," *Engadget*, 13-Jul-2010. [Online]. Available: http://www.engadget.com/2010/07/13/the-joystiq-indie-pitch-0-a-d/. [Accessed: 02-Dec-2015].

57. "0 A.D. (PC)," *IGN*. [Online]. Available: http://www.ign.com/games/0-ad/pc-664227. [Accessed: 02-Dec-2015].

58. R. Wanat and R. K. Mantiuk, "A comparison of night vision simulation methods for video," 2014, pp. 1–8.

59. "File:Gallic Fields.jpg," *Wikipedia, the free encyclopedia*. [Online]. Available: https://commons.wikimedia.org/wiki/File:Gallic_Fields.jpg. [Accessed: 29-Jul-2016].