What Drives People: Creating Engagement Profiles of Players from Game Log Data

Erik Harpstead
Carnegie Mellon University, PA
eharpste@cs.cmu.edu

Thomas Zimmermann, Nachiappan Nagappan
Microsoft Research, WA

Jose J. Guajardo
Microsoft Studios User Research, WA

Ryan Cooper, Tyson Solberg, Dan Greenawalt
Turn 10 Studios, WA

{tzimmer, nachin, jjg, ryanc, tysonsol, dangre}@microsoft.com

ABSTRACT
A central interest of game designers and game user researchers is to understand why players enjoy their games. While a number of researchers have explored player enjoyment in general, few have talked about methods for enabling designers to understand the players of their specific game. In this paper we explore the creation of engagement profiles of game players based on log data. These profiles take into account the different ways that players engage with the game and highlight patterns associated with active play. We demonstrate our approach by performing a descriptive analysis of the game *Forza Motorsport 5* using data from a sample of 1.2 million users of the game and discuss the implications of our findings.

Author Keywords
Player profiles; Engagement; Telemetry

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; K.8.0 Personal Computing: Games.

INTRODUCTION
Understanding who is playing a game is a core value of both game design and game user research. The more designers can understand about their players' preferences and desires the better they can craft a satisfying experience. The HCI literature provides a number of explorations into player enjoyment [25,31,33], and user typologies [7,9,32]. What is less clear is how practitioners can apply the body of existing theory on players' motivation to answer the design questions of their own games.

The nature of player experience is complex and multi-dimensional. A growing body of work in HCI has emerged to better understand this space [25,27]. While more work is exploring the space of why people play games in general, there is a need for techniques that designers can apply at the specific level of individual games. Allowing designers to frame questions in the terms of their own games rather than the language of an abstract model can provide them a better perspective to reason through design alternatives.

CHI PLAY 2015, October 03-07, 2015, London, United Kingdom

ISBN 978-1-4503-3466-2/15/10…$15.00
DOI: http://dx.doi.org/10.1145/2793107.2793114

Figure 1. A screenshot from *Forza Motorsport 5*.

In this paper we demonstrate a technique for creating engagement profiles by segmenting a player population and explore what patterns lead to active and sustained engagement within each segment. This descriptive analysis is based purely on logged game metrics. We make use of log data from a sample of 1.2 million players of the game *Forza Motorsport 5* (Figure 1) a racing game on the Xbox One console.

The contributions of this paper are:

1. A technique for creating engagement profiles of game players based purely on log data.
2. A model for applying existing HCI and game design theories to specific game contexts.
3. A case study creating engagement profiles from a sample of 1.2 million *Forza Motorsport 5* players.

PLAYER SEGMENTATION
The notion that a population of users cannot be accurately described by a single monolithic set of interests is not a new one to HCI. A number of researchers in human-computer interaction and game user research have explored the development of methods for parsing out the nuances present within large groups of users [7,9]. Each of these methods uses different forms of data and is useful for generating different kinds of insights about users.

In the user centered design tradition, the creation of user personas is a common technique for giving designers a personal sense of who uses their products [9]. Personas help designers develop a detailed understanding of potential users when thinking through how a system should behave. A number of researchers have explored the creation and use of personas in the design process [7,15,24]. While they provide designers with a rich cast of characters to relate to in their thinking, some have debated if the actual utility of personas justifies the relative cost of their creation [15].

The field of game studies has a number of player typologies that seek to explain different players' motivations when approaching games [1,2,32]. Tuunanen and Hamari have found that many of the existing player typologies include elements of achievement, exploration, sociability, aggression, and immersion [32]. While these theoretical typologies are useful to the development of a broader theory of game play, they can be hard to apply to games where certain constructs do not exist. For example, Bartle's taxonomy of player types, one of the most famous, was developed primarily in the context of Multi-User Dungeon Role-Player Games [2].

There has been some recent encouraging work in machine learning and artificial intelligence approaches to diagnose the play styles of different individuals. Higgs and colleagues explored the use of Markov chains through game states in an iOS game to see how different players' behaviors mapped to good or bad strategies [17]. In a similar approach Holmgård and colleagues used agent based modeling to see how well players' behavior traces matched to a set of artificial play traces based on designer created strategy agents [18]. While these approaches show much promise, Holmgård and colleagues note, in their current state they require a significant investment in AI development and highly detailed data making them hard to scale to larger games.

Drachen and colleagues are the closest example to our current work that exists in the literature [12]. Their approach makes use of Emergent Self-Organizing Maps, a form of unsupervised learning, to group players according to commonality of in-game metrics. By observing the patterns that arise in the resulting clusters of data they were able to determine four distinct player types in the game *Tomb Raider Underworld*. While this work is useful for diagnosing styles of play, i.e. how players approach the game differently, it does not provide insights into why players play differently.

Outside of the broader HCI literature, there have been a few explorations into ways of segmenting user populations. In the marketing tradition segmentation has commonly focused on demographic based distinction, such as age gender and location [4]. Bhatnagar and Ghose explored the segmentation of e-shoppers based on a model of their purchase probability of different items arising from users' perception of product benefits [4].

With a few notable exceptions [12,17,18] most existing user segmentation approaches involve the use of survey methodologies. While these techniques are certainly useful in that they provide a lens onto users' own self perceptions they are also expensive to produce and can normally only capture a portion of the broader user community. This presents a problem in games with larger player bases. The advent of large scale telemetry solutions in games [13,22,30] makes it easy to get access to user information cheaply, unobtrusively and accurately on large scales.

PLAYER ENGAGEMENT

Player engagement has been studied by many researchers under many different names. A recent review by Boyle *et al*. explores many of the different perspectives researchers have applied to engagement in the past [6]. Using their categorization our work would fit into the space of engagement as game usage, however, unlike prior work in that space we are interested in looking at patterns of actual use rather than self-reported use.

A highly relevant exploration of player engagement was done by Theng *et al.* who looked at a number of possible factors of players' motivations in playing games [31]. What is useful about this work is the consolidation of a number of models of general user motivations with respect to the particular interest of games. Their resultant Playability Acceptance Adoption Model (PAAM) resembles the Technology Acceptance Model (TAM) [10] where perceived ease of use and perceived benefits both contribute to a player's satisfaction with a game which in turn leads to continued play. The main difference between PAAM and TAM is what factors mitigate the perceived ease of use and perceived benefits, where PAAM includes a number of game design elements. One of the drawbacks of Theng's study is that while the outcome of their model is continued play, the study itself was based on general survey data and could only approximate intention to continue playing rather than actual play.

Another relevant exploration of player engagement comes from Bouvier and colleagues [5]. They combined Self-Determination Theory (SDT) [29], Activity Theory (ADT) [20] and Trace Theory [8] to qualify users' engaged behaviors in a social game. In their work they were interested in understanding how engaged behaviors are different from unengaged ones on an interaction trace level. They separated engagement into four categories: environmental, social, self, and action. Their approach involves the identification of engaged user traces and a detection scheme for finding such traces among general player traces. Their method was capable of reaching reasonable agreement with human raters when judging the engagement of 12 players. While the approach and theoretical backing are promising the fine grain level of detail required to perform their analyses is prohibitive for larger games.

A more design theoretic approach to player engagement can be seen in Hunicke, LeBlanc, and Zubek's Mechanics, Dynamics, and Aesthetics (MDA) model [19]. MDA breaks a game's design into three components: mechanics, which are the essential rules or systems of a game; aesthetics, which are the core engagements of the game and the feelings it elicits in the player; and dynamics, which are the emergent interactions between the player and the game while it is being played. In this model players are drawn to different kinds of aesthetics, e.g. challenge or fellowship, based upon their preferences.

Our approach seeks to draw on the prior bodies of work on player segmentation and player engagement to define player engagement profiles. These profiles take into account differential patterns of play, as informed by segmentation research. We then employ prior engagement and motivation research to understand what game features appear to drive engagement with the game within groups. We do this by creating user groups based on player behavior patterns and examine the statistical relationships between users' engagement and reward functions present in the game. Given these models of players, designers can consider what elements of their game are important to their highly engaged users when considering what to do during design iteration.

FORZA MOTORSPORT 5

Forza Motorsport 5 (FM5) is a racing game developed for the Xbox One by Turn 10 Studios. It is the fifth installment of the Forza franchise as well as the first to appear on the Xbox One. Released as a launch title for the Xbox One platform in November of 2013 the game has become the fastest selling racing game in Xbox history with more than one-third of console owners purchasing the game since its November launch [34].

Core tenants of the Forza franchise are an emphasis on realism and an expression of car culture. FM5 provides players the ability to drive a number of real world cars, ranging from every day commuter cars to exotic super cars, on famous real world tracks. Beyond racing, players also have the ability to customize the visual aspects of their cars with unique paintjobs and livery designs. FM5 also allow players to tune the performance of their cars to optimize for preferred power, handling, and driving style.

Racing is the primary form of interaction in FM5 and there are a number of different race modes that players can choose between, each with their own particular rule sets.

The eight game modes in FM5 include:

- *Career* - players progress through a number of different leagues each centering on a particular class of car. Career is the primary single player mode in the game and one of the primary catalysts for acquiring new cars.
- *Multiplayer Hopper* - players race online against other human racers who are selected randomly from a public hopper optimizing for roughly equal skill.
- *Multiplayer Private* - players setup their own private races against friends or others they invite. Players can impose any restrictions on these races that they want.
- *Free Play* - players engage in a number of different open format races. This mode also provides opportunities for players to race in cars they do not yet own in the game.
- *Rivals* - players race against the recorded lap times of other racers. A global leaderboard of recorded times is maintained for each track in the game.
- *Test Drive* - players race on a track alone to try out new car configurations or practice sections of a track. This mode is part of the testing loop when developing new car configurations.
- *Split Screen* – players race with each other on the same system sharing a screen. This is the local multiplayer option.

The Forza franchise is over 10 years old and as the series continues to develop into the future it is important to continually take stock of what the player community values. The structure of FM5 provides players with a number of different ways to engage and it would be helpful to better understand which ones players seem to engage with the most and what elements drive that engagement.

OUR APPROACH

Our approach seeks to provide insights into the composition of a games' player base as well as what features appear to drive their engagement. While we take inspiration from existing models, such as PAAM [31] and MDA [19], we recognize that it is important to understand players in terms of the game in question. A key interest of the approach is to remain grounded in the specific game rather than map the game directly to a general theory of game motivation or engagement.

The first assumption our method makes, following from player typology research, is that players will fall into distinct behavioral groups based on what features of the game they interact with. This differential use is particularly focused on different modal forms of play. Second we assume that the people in different behavioral groups not only play differently but do so for different reasons. To get an understanding of this component we borrow from the theory of the PAAM model to find the game features that are capable of predicting players' engagement with the game. This focus on actually observed play is crucial because it is more informative in understanding the state of a current design.

Applying this structure to a particular game requires the definition of three categories of metrics (Figure 2). While the particular metrics we use in our analysis are contextual to FM5 we feel it is reasonable to assume that all games will have analogous concepts in their designs.

Three Metric Categories

We refer to the first type of metric as a **behavioral** or interaction metric. These metrics are measures of ways in which players interact directly with the game. They might represent numbers of sessions in particular game modes or paths taken through the game space. The primary use of these behavioral metrics is to group players who interact with the game in similar ways in order to produce behavioral segmentation. In our work on FM5 we use the percentage of races players spent in each of the different game modes as behavioral metrics.

The second type of metric is **engagement** or retention. These metrics should be familiar in most contexts as measures of sustained use. This is similar to metrics used in churn analyses [16], however, we are more interested in understanding the behaviors of players who are still playing and have not churned. In our context we explore the use of 2 different measures of engagement:

1. Whether a player has logged a race in the past month (30 days). This is a measure of active engagement as it denotes someone who is part of the currently active player base.
2. The lifetime total number of races each player has raced. This serves as an overall measure of engagement intensity.

The third type of metric is what we call a **reward** metric. These metrics represent elements of the game which players could possible care about or derive value from. In defining them we are interested in balancing both elements of experience that prior theory would say could be rewarding to players as well as take inspiration from the existing design canon of a franchise. Reward metrics generally manifest as forms of feedback to the player but proxies can be develop when specific feedback is not available. Examples could include collecting items, beating certain kinds of opponents, or completing quests.

In our work we mostly took inspiration from the 8 aesthetics listed within the MDA model [19]. Our goal was to find a particular metric that could be mapped to each aesthetic in a way that could differentially experience by players. Sensation, for example, was not considered appropriate because all players have the same visual experience with the game. In places where there was conflict between how to align an aesthetic to a single metric, e.g. challenge could be conceived in multiple ways, we turned to the design canon of the game to inform how an aesthetic could be interpreted. The rewards metrics that we used in our examination of FM5 are:

Podium Percentage (PP) is measured as the life time percentage of races where the player podiumed, i.e. received a medal. Every race offers the chance for the player to receive a medal (bronze, silver or gold) based on the particular event type of the race. In a standard race format medals are award based on finish position. In other formats, medals might be given out for other criteria like passing a certain number of cars or completing a lap in a specified amount of time. Since this metric is represents direct skill at the primary task in the game it is related to a challenge aesthetic as well as a mastery motivation [29].

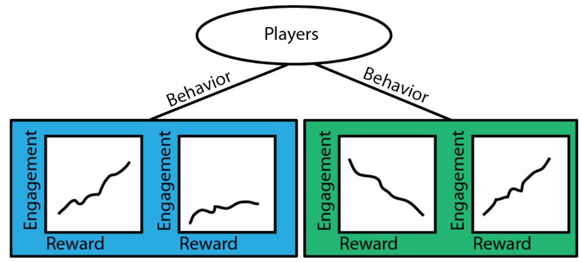

Figure 2. Engagement Profiles arises from what rewards engage players within behavioral groups

Time Ratio Mean (TRM) is measured as the life time average of how fast the player drives relative to the fastest possible AI racer for a given track; where a lower score is better, and a negative score implying a player is capable of beating the best possible AI. The best possible AI time is a known feature of each track as part of FM5's drivatar AI system [26]. This metric is also related to an aesthetic of challenge and a mastery motivation but focuses more on the dimension of speed as central to the racing experience rather than winning, which some players may value more.

Replay Rate (RR) is the percentage of races players initiated with the same car, track, and game mode as their previous race. This approximates a measure of grinding and persistence where players engage in the same experience repeatedly. This metric is inspired by what the MDA model refers to as an aesthetic of submission or abnegation where players engage with a game as a mere pastime [19].

Friend Count (FC) is the total number of friends the player has associated through their Xbox Live account. A number of existing game motivation theories highlights the importance of social dimensions of the player experience [19,33]. The number of friends a player has allows us to approximate the effect of social interaction on engagement.

Livery Use (LU) is the percentage of races where the player uses a custom livery on their car, giving it a different visual appearance. FM5 provides a rich toolset for creating and sharing custom liveries for cars as a form of personal expression. This usage percentage is whether players used a livery in a race. It does not capture if the player designed the livery themselves. Regardless of who created the original design, the use of a livery, instead of a basic visual design, is a metric of self-expression, which is also regarded as an aesthetic in the MDA model [19].

Tune Rate (TR) is a measure of how often, on average, the player tuned their car. Tuning is a unique aspect of the car culture surrounding FM5 and involves meticulously tweaking the configuration of a car for optimal racing performance. This metric is calculated by taking the number of unique tuning configurations seen in a player's log and dividing by the number of unique cars. This metric is associated with both a mastery motivation and a self-expression motivation [19,29]. It also represents a central aspect of the FM5 experience and so is likely to be important to players.

Realism Use (RU) is a measure of what percentage of races players used realistic race settings. FM5 prides itself on delivering realistic racing experiences. To facilitate this the game provides players a number of options in how to craft their preferred experience [11]. At the extreme end of these options is the simulation mode where players race in full realistic conditions requiring pit stops to refuel and cope with damage to their car. This metric is associated with both a fantasy and sensation aesthetic [19] as well as being tied to a core part of the FM5 experience.

Car Class (CC) is a measure of the average performance level of cars used by the player across all races. All cars in FM5 fit into one of 8 classes, which rank them on features like power and handling. Higher classes represent the more powerful and exotic vehicles where lower classes represent more common street cars. As the fantasy element of being able to drive rare cars is another core engagement of the Forza franchise we include this metric to approximate players' desire for driving fantasy.

We view each of these metrics as a design hypothesis of what engages people to play FM5. By this we mean that according to existing theories of game motivation and engagement and the established design knowledge of the Forza franchise, players who are highly engaged with FM5 will show an affinity for some mixture of these metrics. Additionally, we do not necessarily expect all of these design hypotheses to hold across each of the behavioral group, rather we expect that some groups will show stronger relationships for some metrics over others.

Method

At a high level, our approach involves the use of a set of behavioral metrics as a schema for grouping players. It then performs a regression of engagement on reward metrics within each behavioral group to determine which rewards contribute most to players' sustained engagement. The result is a profile of rewards that describes the factors that motivate the players within each behavioral group.

There are number of methods for generating groupings of users from quantitative data [4,12]. Because we have a very large set of players we employ the Clustering for Large Applications (CLARA) algorithm [21], which employs a sampling strategy to perform k-medoids on a large set of data. We use the silhouette method for determining the appropriate number of clusters k to form, choosing the k within a range which produces clusters with the maximum average silhouette width [21].

Once an appropriate grouping of users is obtained we perform a regression analysis within each group. This regression looks to explain how much each of the reward metrics contributes to users' engagement metrics within a behavioral group. To aid in interpretation regressions are performed in terms of standard deviations of the reward functions rather than their raw values.

The distribution of response variables is an important aspect to consider when selecting which kind of regression to perform. There has been some discussion of which form of distribution is most appropriate for player life time usage in games [3]. A common mistake is to assume that data are normally distributed and fit a standard linear model, when this is not always the case in practice. In our data the classification of whether or not a player has played recently is a binary response variable and so it is appropriate to perform a logistic regression [23]. The engagement intensity metric is a total number of races done by a player and so represents count data, which is best fit with a Poisson regression [23].

RESULTS

The results we describe here are derived from a sample of the FM5 telemetry logs spanning from the release of the game (Nov. 22, 2013) for a 200 day period (June 10, 2014). This sample contains 120 million race entries from a sample of 1.2 million players.

In this sample the engagement intensity metric in the sample is the total number of race logs provided for each player. The metric of active engagement is determined by whether or not a player had logged a race in the 30 day period leading up to June 10, 2014. It is worth noting that while a 30 day cut-off is intuitively attractive, as it roughly equates to a month, it is an imperfect measure of active engagement. In the sample 47% of the players treated as actively engaged could have been considered inactive at one point in their player lifetime, i.e. there is a gap between races of more than 30 days at some point in their player lifetime. This means that this metric *under estimates* active engagement.

Clustering

We use race entry data to determine what percentage of all races recorded for each user were spent in each of the different game modes. These relative percentages were employed as the behavioral metric forming the basis for our clustering. A Duda-Hart test determined that there was more than 1 cluster present in the data so we proceeded to run the CLARA algorithm allowing k to range from 2 to 40. After this search process 10 clusters of players were found in the data. The mean time spent in each game mode for each group along with the overall population is plotted in

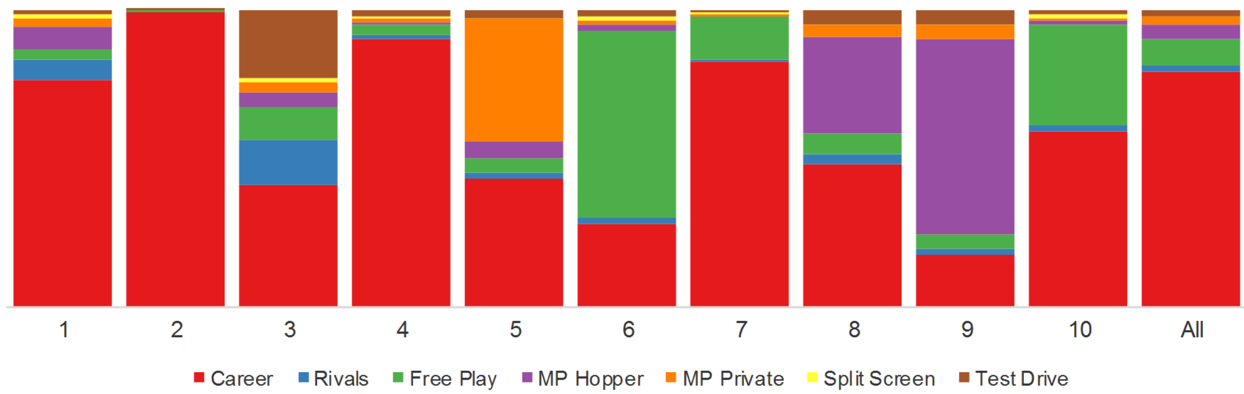

Figure 3. A plot showing the average percentage of time spent in each game mode by players in each behavioral group as well as in the overall population.

Figure 3. Table 1 shows each group's share of the overall sample and of the subset of players raced recently, i.e. within 30 days of the sample date. Table 1 also shows the mean race count within each group normalized for confidentiality.

A few interesting patterns are visible in the race counts and game mode percentages. The first noticeable trend is that Career mode appears to be the most popular mode in the overall population and in almost every group. This is not unexpected because the game starts out guiding players through their first career series, consisting of 10 races, before setting them free to try other modes.

Another high level pattern that was not readily expected is the comparatively low rate of use for the split screen mode across all the groups. While users commonly express a desire for a local multiplayer option their actual usage patterns seem to contradict this sentiment.

An interesting pattern visible in Table 1 is that the two smallest groups (3 and 9) also have the highest mean race counts. In fact while groups 3 and 9 represent only 6.18% of the overall population, they account for an average 21.65% of all races.

Group	Percentage of Overall Sample	Percentage of Recent Players	Mean Race Count
1	9.27%	9.27%	2.8X
2	38.09%	12.95%	X
3	3.35%	6.38%	6.0X
4	13.62%	10.77%	1.9X
5	3.78%	4.60%	3.9X
6	3.98%	5.30%	2.3X
7	12.43%	9.83%	1.4X
8	4.77%	6.95%	5.0X
9	2.83%	5.15%	9.3X
10	7.88%	7.93%	1.6X

Table 1. Descriptive statistics of engagement metrics within each group.

Turning to the composition of the individual groups themselves, the first cluster that stands out is Group 2 who spend nearly 100% of their time in the career mode, spending only a combined average of 1.2% of races in any other mode. This group also accounts for a smaller portion of recent players compared to their share of the overall population, suggesting that they are mainly people who tried out the game but have not been playing recently. This kind of population is not unusual for games of this size, particularly for launch titles of a new gaming platform. Exploring what factors contribute to being actively engaged in this group can help in understanding what could get these players playing more often.

Groups 6 and 9 also stand out as being primarily concerned with one specific mode, free play and multiplayer hopper respectively. Each of these groups clearly has a core engagement in the game that is not the career mode.

Groups 5, 8, and 10 also have a solid interest in one other mode that is not career but still spend a large amount of time in career mode. Group 5 in particular is interesting for their frequent use of the private multiplayer mode, which is generally rare among all of the other groups.

Groups 1, 4, and 7 all appear to primarily play career mode while dabbling in other modes, with the main distinction between groups being what they dabble in. There are also some differences among the groups in terms of race count but their general levels of engagement are relatively close. If a smaller set of sub-populations is desired by designers for the purposes of considering their population, merging these three groups could be entertained as a possibility.

Group 3 at first glance appears to be a group with no real core engagement in the game, instead sampling among the different modes. Upon further consultation with the game designers it appears that these players may be "hot lappers," who are people that primarily focus on getting faster lap times in order to climb the global leaderboards. The rivals mode is where players race against other players' recorded lap times on the global leaderboard to attempt to raise their rank while the test drive mode is commonly used to try out new car configurations and practice difficult portions of tracks. This “hot lapper” interpretation was reinforced by looking at the regression data.

Regression

Within each group and for the overall population we performed 2 regressions. The first was a logistic regression of all reward metrics on whether a player had raced recently. The second was a Poisson regression of the reward metrics on race count. Both regressions were performed with reward metrics centered to their means and scaled by their standard deviations in order to consider the relative contribution of each metric on the same scale (descriptive statistics for each of the reward metrics are shown in Table 2). The results of these regressions can be seen in Table 3 (for recency) and Table 4 (for race count).

When interpreting logistic regression models, the coefficient β_i corresponds to the change in the log of the odds for a one unit, in our case standard deviation, change in factor x_i given that all other factors remain fixed. The change of odds ratio (*OR*), i.e. the change in percentage chance a player logged a race in the past 30 days, for a standard deviation change of factor x_1 can be computed by raising *e* to the power of the logistic coefficient, $OR = e^{\beta_1}$.

Interpreting Poisson regression models follows a similar pattern with the exception that the coefficient β_i corresponds to the change in the log of the expected count (*EC*), i.e. log of total race count, for a one unit change in factor x_i given all other factors remain fixed. The ratio of change in the expected count for a standard deviation increase of factor x_1 can be computed by raising *e* to the power of the Poisson coefficient, or $EC = e^{\beta 1}$.

Given that statistically significant differences are easier to detect in large data sets we use an alpha level of .001 to determine if a relationship is significant. Additionally, we apply a Bonferroni adjustment to control for the number of statistical tests and avoid the random chance of finding significance [14]. To do this we divide the starting alpha level (.001) by the number of statistical tests being performed, (1 + the number of rewards) × (1 + the number of groups) × 2, resulting in a final alpha value of $5.05\mathrm{E}^{-6}$. If a reward’s coefficient’s p-value is still less than the adjusted alpha value then the coefficient is considered significantly different from 0 meaning the reward is likely contributing to the explanation of player engagement. Before performing each regression we assessed the multicollinearity between predictor variables using the variance inflation factor [28] and found them all to be within an acceptable range.

While the significance of regression coefficients tells us which rewards are likely explanations of player engagement, what is more interesting is which ones among the likely motivations are the strongest. Since all of the reward coefficients are in the scale of each reward’s standard deviation, they can be compared in terms of their magnitudes with higher magnitudes indicated stronger effects from a given reward. In both Table 3 and Table 4, coefficients that result in at least a 25% change in the response variable have been emphasized as strong effects.

Metric	Mean	Std. Dev.	Min	Max
PP (Podium %)	0.76	0.24	0.00	1.00
TRM (Time Ratio Mean)	0.09	0.06	-0.23	0.93
RR (Replay Rate)	0.09	0.11	0.00	1.00
FC (Friend Count)	29.51	45.81	0.00	7066.00
LU (Livery Use)	0.80	0.21	0.00	1.00
TR (Tune Rate)	1.37	0.69	1.00	129.17
RU (Realism Use)	0.15	0.30	0.00	1.00
CC (Car Class)	3.18	0.82	1.01	7.97

Table 2. Descriptive statistics of the reward metrics in the overall population.

Looking at the results of the logistic regression of active player engagement (Table 3) it is apparent, both across groups and within the overall population itself, that the Livery Use (LU) and Car Class (CC) appear to be the strongest explanatory variables. This would mean that players who more commonly customize their car’s appearance and drive higher performing cars are more likely to continue to be active with the game.

A somewhat weaker pattern can be seen across groups in the Realism Use (RU) dimension. The RU pattern shows strong effects for groups 1, 5, 8 and 9 meaning that players in these groups who raced with realistic mode were more likely to remain actively engaged. Interestingly, these groups are also the three whose core engagement with the game is a multiplayer mode.

The Friend Count (FC) metric has a strong relationship to engagement in Groups 3, 5, and 9. Groups 5 and 9 both have strong core engagements in multiplayer modes, which would connect well with a social dimension of motivation. It is less obvious why Group 3 would have an effect from having more friends but the “hot lapper” interpretation of this group points to a strong competitive element where players are competing often with each other through the global leaderboard.

There is an interesting pattern looking at the Podium Percentage (PP) reward across groups. Groups 1, 2 and 9 all seem to be somewhat positively engaged by medaling in races. Given that Group 9 is strongly focused on public multiplayer it makes sense that they would be more engaged by placing in races. The patterns in Groups 1 and 2

Group	PP (Podium %)	TRM (Time Ratio Mean)	RR (Replay Rate)	FC (Friend Count)	LU (Livery Use)	TR (Tune Rate)	RU (Realism Use)	CC (Car Class)
1	0.112	-0.213	NS	0.121	**0.408**	0.091	**0.240**	**0.440**
2	0.100	-0.167	NS	0.052	**0.756**	**0.283**	0.209	**0.367**
3	-0.123	-0.246	-0.187	**0.235**	**0.556**	0.047	0.219	**0.336**
4	NS	-0.223	NS	0.084	**0.369**	NS	0.181	**0.394**
5	-0.158	-0.267	-0.068	**0.227**	**0.441**	0.076	**0.320**	**0.333**
6	-0.162	NS	-0.269	0.088	**0.312**	0.150	0.162	**0.235**
7	NS	-0.159	NS	0.079	**0.370**	0.118	0.178	**0.389**
8	NS	-0.211	NS	0.193	**0.322**	0.103	**0.304**	**0.378**
9	**0.246**	-0.115	-0.055	**0.267**	**0.360**	0.090	**0.468**	**0.247**
10	NS	-0.118	-0.170	0.085	**0.408**	0.157	0.184	**0.311**
Overall	-0.159	-0.285	-0.027	0.161	**0.413**	0.195	**0.240**	**0.480**

Table 3. Regression coefficients for the logistic regression of players' recent engagement on reward metrics within behavioral groups. Bolded coefficients indicate effects that result in at least a 25% change in log odds of engagement per unit of reward.

seem to indicate those that can continue to progress through the game are more likely to remain playing. On the other side of the PP reward, Groups 3, 5, and 6 seem to be negatively affected by winning races. Given that Group 6 is primarily interested in the free play mode, it follows that they would be less focused on winning races; Group 5 is strongly engaged in racing online with friends, which could mean they are less driven by winning; and "hot lapper" interpretation of Group 3 supports an interest in lap times without necessarily caring about winning a race.

Group 2 also has a moderately strong effect from the Tuning Rate (TR). This would mean that people in the group who have done some experimenting with tuning their car are more likely to remain actively engaged. As Group 2 is generally a smaller portion of the actively engaged population, this is useful information when considering what they might find interesting.

Interpreting the results in Table 4 is similar but has nuanced differences from Table 3. The high level patterns visible across groups and in the overall population highlight Livery Use (LU) and Time Ratio Mean (TRM). This echoes the importance of self-expression from the previous regression while also showing that players who are faster tend toward racing more, recall that a lower value of Time Ratio Mean indicates a better time relative to optimal.

The Car Class (CC) metric is strong for Groups 1, 2, 4, 7, and 10. This is interesting because those four groups also have generally low mean race count (Table 1). This would suggest that among the groups with lower engagement intensity driving more powerful cars and fulfilling the fantasy of FM5 tends to lead to more intense engagement.

Group	PP (Podium %)	TRM (Time Ratio Mean)	RR (Replay Rate)	FC (Friend Count)	LU (Livery Use)	TR (Tune Rate)	RU (Realism Use)	CC (Car Class)
1	**0.532**	**-0.305**	0.196	0.022	**0.510**	0.086	0.097	**0.321**
2	**0.286**	**-0.394**	0.090	-0.065	**0.637**	**0.225**	0.105	**0.418**
3	0.004	**-0.358**	0.147	0.044	**0.759**	0.010	0.073	0.086
4	**0.341**	**-0.359**	0.143	-0.008	**0.325**	0.097	0.081	**0.344**
5	0.115	**-0.382**	0.210	0.088	**0.662**	0.050	0.134	0.165
6	-0.031	-0.180	0.077	0.022	**0.462**	0.020	0.169	0.160
7	**0.320**	-0.283	0.138	-0.012	**0.370**	**0.227**	0.115	**0.337**
8	**0.389**	**-0.351**	0.151	0.075	**0.665**	0.077	0.086	0.152
9	**0.346**	-0.228	0.131	0.037	**1.073**	0.017	0.051	0.039
10	0.179	-0.239	0.025	0.017	**0.517**	0.122	0.161	**0.258**
Overall	-0.041	**-0.539**	0.120	0.038	**0.671**	0.010	0.097	**0.313**

Table 4. Regression coefficients for the Poisson regress of players' total race counts on reward metrics within behavioral groups. Bolded coefficients indicate effects that result in at least a 25% change in expected race count per unit of reward.

The race count regression results provide further support for the “hot lapper” interpretation of Group 3 where it is clear that winning races (PP) has effectively no noticeable contribution to their continuing to race. Conversely improving in speed does have a positive impact on continued play.

One of the most interesting patterns across both regressions is that Livery Use (LU) has such a strong effect on engagement. The base rate of livery use is high (Table 2) because players are prompted to select a livery design every time they buy a new car. This means that not using a livery represents a conscious choice by the user. The regression results then suggest that those who actively opt out of using a livery are less likely to be engaged.

DISCUSSION

In this paper we described a method for generating engagement profiles of players based on a set of three types of metrics. Players are first grouped based on behavioral similarity then within the groups their engagement is explained by a number of reward functions to consider which rewards are the strongest for different groups of players.

One of the major benefits of this approach is its relatively light application cost, provided that a game design team is already capturing a base of telemetry data. Clustering and regression algorithms are commonly available in most statistical packages. The main cost of application is in framing behavioral, reward, and engagement metrics in terms of the game in question.

Another place where the approach excels is in established games and franchises. The original *Forza Motorsport* was released in 2005, making the franchise over a decade old. The continued evolution of the franchise has generated a deep pool of design knowledge specific to the Forza world. By developing a technique that demands little in terms of theoretical re-framing this established design knowledge can be brought to bear in the analysis.

When developing reward metrics there are two general approaches one might explore. The first approach is to have designers consider what elements of the game they expect users to be attracted to and develop reward metrics which approximate those theories. In this framing, designers’ intuitions represent design hypotheses about player motivations that can be confirmed or rejected by player data. This is the approach we took in our current analysis.

The second approach to metric creation is to survey users to determine what kinds of things they enjoy doing in the game. In this case our approach is best suited as a means of seeing how common certain player attitudes are in the broader population. It can also be employed in a multi-tiered approach as a way of scaling up survey concerns to a broader population of users.

While we believe this approach is a useful means for checking design intuitions on players’ against real data we feel it is prudent to recognize its limitations. When interpreting the regression results it is important to remember that the final conclusion returned is the best explanation among those provided. It is always possible, as in all statistics, that another unmeasured variable could better explain player engagement. Also, any change to the list of candidate reward metrics will alter the results for all metrics requiring a reinterpretation of results.

When developing a collection of metrics it is important to consider the relationships between variables before assigning them to any one of the three metric categories. For example, early on in our own exploratory data analysis of FM5 we considered using a count of cars owned as a metric meant to capture people wanting to collect content however players who have played more have necessarily purchased more cars, because of how the game is designed. Given this direct relationship the regression process would have been confounded by the underlying relationship of these variables.

Another potential issue with the approach is that the interpretation of the results is a purely behavioral claim. There is no guarantee that players would self-identify with these player profiles if asked directly. While we acknowledge this limitation we would argue that the intent of the method is to get an understanding of the general user trends that exist in a player base at a high level rather than capture players’ self-professed tastes. Taking this broader view it is understandable that some of the texture of individual users would be lost in the process.

A final drawback, which also highlights a potential for future work, is that the final product of user profiles is communicated as a series of regression coefficient tables. These tables require statistical training to properly interpret, particularly when non-standard generalized linear models are used. Developing ways of better communicating such results would further benefit designers who might not have the data science expertise to form their own conclusions.

CONCLUSION

Understanding who players are and what about a game they value is a crucial component of game design. We have demonstrated a way of conceptualizing a broad base of players in terms of their behavior and response to different in game reward functions. This approach has helped us greatly in developing a picture of the players of FM5. We hope others can find value in applying a similar approach to their work with their own games.

ACKNOWLEDGMENTS

Erik Harpstead performed this work as an intern at Microsoft Research. We thank Sauvik Das for his feedback and the staff at Turn 10 Studios for their work on *Forza Motorsport 5*.

REFERENCES

1. Alix, A.Beyond P-1: Who Plays Online? *Proc. DiGRA 2005*, (2005).
2. Bartle, R.Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD Research 1*, 1 (1996), 19.
3. Bauckhage, C., Kersting, K., Sifa, R., Thurau, C., Drachen, A., and Canossa, A.How Players Lose Interest in Playing a Game : An Empirical Study Based on Distributions of Total Playing Times. *Proc. CIG 2012*, IEEE (2012), 139–146.
4. Bhatnagar, A. and Ghose, S.A latent class segmentation analysis of e-shoppers. *Journal of Business Research 57*, 7 (2004), 758–767.
5. Bouvier, P., Sehaba, K., and Lavoué, É.A trace-based approach to identifying users' engagement and qualifying their engaged-behaviours in interactive systems: application to a social game. *User Modeling and User-Adapted Interaction 24*, 5 (2014), 413–451.
6. Boyle, E. a., Connolly, T.M., Hainey, T., and Boyle, J.M.Engagement in digital entertainment games: A systematic review. *Computers in Human Behavior 28*, 3 (2012), 771–780.
7. Canossa, A. and Drachen, A.Patterns of Play: Play-Personas in User-Centred Game Development. *Proc. DiGRA 2009*, (2009).
8. Clauzel, D., Sehaba, K., and Prié, Y.Enhancing synchronous collaboration by using interactive visualisation of modelled traces. *Simul.Model. Pract. Theory 19*, 1 (2011), 84–97.
9. Cooper, A.*The Inmates Are Running the Asylum: Why high-tech products drive us crazy and how to restore sanity.* Sams, 1999.
10. Davis, F.D., Bagozzi, R.P., and Warshaw, P.R.User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science 35*, 8 (1989), 982–1003.
11. Debeauvais, T., Zimmermann, T., Nagappan, N., et al.Off With Their Assists: An Empirical Study of Driving Skill in Forza Motorsport 4. *Proc. FDG 2014*, (2014).
12. Drachen, A., Canossa, A., and Yannakakis, G.N.Player modeling using self-organization in Tomb Raider: Underworld. *Proc. CIG 2009*, IEEE (2009), 1–8.
13. Drachen, A. and Canossa, A.Towards gameplay analysis via gameplay metrics. *Proc. MindTrek 2009*, ACM Press (2009), 202.
14. Dunn, O.J.Multiple Comparisons Among Means. *Journal of the American Statistical Association 56*, 293 (2014), 52–64.
15. Friess, E.Personas and Decision Making in the Design Process: An Ethnographic Case Study. *Proc. CHI 2012*, ACM Press (2012), 1209–1218.
16. Hadiji, F., Sifa, R., Drachen, A., Thurau, C., Kersting, K., and Bauckhage, C.Predicting Player Churn in the Wild. *Proc. CIG 2014*, IEEE (2014), 1–8.
17. Higgs, M., Girolami, M., Morrison, A., and Chalmers, M.Analysing User Behaviour Through Dynamic Population Models. *Ext. Abstracts CHI 2013*, ACM Press (2013), 271–276.
18. Holmgård, C., Liapis, A., Togelius, J., and Yannakakis, G.N.Generative Agents for Player Decision Modeling in Games. *Proc. FDG 2014*, (2014).
19. Hunicke, R., Leblanc, M., and Zubek, R.MDA : A Formal Approach to Game Design and Game Research. *Proc. of the AAAI Workshop on Challenges in Game AI*, (2004), 1–5.
20. Kaptelinin, V. and Nardi, B.A.*Acting with Technology: Activity Theory and Interaction Design.* MIT Press, Cambridge, MA, 2006.
21. Kaufman, L. and Rousseeuw, P.J.*Finding Groups in Data: An Introduction to Cluster Analysis.* Wiley, 2009.
22. Kim, J.H., Gunn, D. V, Schuh, E., Phillips, B.C., Pagulayan, R.J., and Wixon, D.Tracking Real-Time User Experience (TRUE): A comprehensive instrumentation solution for complex systems. *Proc. CHI 2008*, ACM Press (2008), 443–451.
23. McCullagh, P. and Nelder, J.A.*Generalized Linear Models.* Chapman and Hall, London, 1989.
24. McGinn, J. (Jen) and Kotamraju, N.Data-driven persona development. *Proc. CHI 2008*, ACM Press (2008), 1521.
25. Mekler, E.D., Bopp, J.A., Tuch, A.N., and Opwis, K.A systematic review of quantitative studies on the enjoyment of digital entertainment games. *Proc. CHI 2014*, ACM Press (2014), 927–936.
26. Microsoft Research.Drivatar™. http://research.microsoft.com/en-us/projects/drivatar/.
27. Nacke, L. and Drachen, A.Towards a Framework of Player Experience Research. *Proc. EPEX 2011*, (2011).
28. O'Brien, R.M.A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity 41*, 5 (2007), 673–690.
29. Ryan, R.M. and Deci, E.Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist 55*, (2000), 68–78.

30. Seif El-Nasr, M., Drachen, A., and Canossa, A., eds.*Game Analytics: Maximizing the Value of Player Data.* Springer, 2013.

31. Theng, Y.-L., Wanzhen, L., Zhiqiang, L., Robert, C., and Pallavi, R.Investigation of the antecedents and consequences of gamer satisfaction. *Computers in Entertainment 9*, 3 (2011), 1–26.

32. Tuunanen, J. and Hamari, J.Meta-synthesis of player typologies. *Proc. Noridc DiGRA 2012*, (2012).

33. Williams, D., Yee, N., and Caplan, S.E.Who plays, how much, and why? Debunking the stereotypical gamer profile. *Journal of Computer-Mediated Communication 13*, 4 (2008), 993–1018.

34. Xbox Wire Staff.Forza Motorsport 5 Free Road America Track Add-On Now Available. 2014.