Understanding Players' Identities and Behavioral Archetypes from Avatar Customization Data

Chong-U Lim
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA
Email: culim@mit.edu

D. Fox Harrell
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA
Email: fox.harrell@mit.edu

***Abstract*—Virtual identities are an integral part of peoples' lives, from online shopping accounts to social networking profiles, from intelligent tutors to videogame avatars. In many videogames, players construct avatars to represent themselves within virtual environments and research has shown that players' sociocultural identities influence their avatar construction and can be a proxy for inferring their values in the non-virtual (real) world. In this paper, we present a computational approach to modeling players' real-world identities using behavioral data collected during the avatar customization process. We used archetypal analysis on player interaction data to develop "behavioral archetypes", which are models of prototypical behavior patterns exhibited by players during the avatar customization process. We modeled patterns of (1) "avatar gender-preferring" behaviors (preferences for a particular avatar gender), (2) "styler" behaviors (preferences for different parts of their avatars, e.g., hair-styler, head-styler, etc.,) and (3) preferences for using avatars of a different gender ("gender-bending") or the same gender ("gender-synchronizing") as the players'. In a user-study with 190 participants, the behavioral archetype model trained via supervised learning had high accuracy (81%) in classifying players' real-world gender using only behavioral data. We show that behavioral archetypes are effective for understanding players in terms of their customization behaviors, real-world genders, and virtual avatar genders.**

I. Introduction

Many videogames provide avatar or character constructors that allow players to select and customize their virtual identity representations. The formation of identity in relation to players and their virtual representations can be characterized by Zach Waggoner's discussion of *avatars* [1]. Rather than just focusing on the technical implementation of avatars, Waggoner calls our attention to the "liminal space between the user and the videogame avatar, between the materiality of the player and the imagination." Indeed, in videogames, avatars are often associated with a player's virtual representation through *characters*, which are more than just proxies for the player, but are imaginatively embedded in the narratives of videogames through various means such as numerical attributes, behavioral characteristics, and backstories. These characters are avatars[1] and, while technically implemented as virtual identities, should not be viewed simply as results of a user-directed creation process [2]. Research has demonstrated that the way players behave in both the real and virtual world can be influenced by these virtual avatars [3], [4]. These avatars and behaviors may also correspond to and reveal aspects of a player's real-world identity such as their gender, race, personality traits, or motivations for play [5], [6]. While user metrics are often deployed to customize gaming experiences for players, preferences do not fall neatly along the lines of race, ethnicity gender, etc. such as those studied via demographics-base methods or self-reported surveys. Our motivation thus lies in developing computational methods that reveal user categories, which emerge from the data and cut across demographics and personality types. This serves social needs of better serving diverse users, commercial needs for expanding the marketplace to better serve specific consumer groups, and expressive needs for tailoring content to users' desires, preferences, and values. Our previous work (outlined in Section III) looked at "infrastructural values" built into systems, such as how the distribution of statistical attributes of characters in games reflected social phenomena potentially symptomatic of bias or world views implicitly shared by the developers themselves. Here, our focus shifts to studying "user values" exhibited by players and that are revealed implicitly from behaviors enacted out within systems.

In this paper, we presents results from *AIRvatar*, our system that collects game telemetry data from avatar creation systems and analyzes them to create computational models of players and their behaviors. In a user-study, 190 participants constructed 16-bit fantasy-styled RPG player characters using an avatar creator we developed. We used archetypal analysis (AA) on telemetry data collected during customization to computationally model the prototypical behavioral patterns of these players, which we term "behavioral archetypes." We evaluated these behavioral archetypes and discovered behavioral patterns of players based on their real-world gender[2] and their preferred choice of avatar gender. We highlight the characteristics of different types of players that were revealed from these behavioral archetypes. We validated our models, constructed using only in-game behavioral data, by their performance in classifying players' real-world genders. To the best of our knowledge, using AA to computationally model avatar customization behaviors to gain insight into players and their identities has not been previously performed.

The rest of the paper is structured as follows: Section II provides an overview of related research and background information on avatars and virtual identities. We cover player modeling, game data mining, archetypal analysis, and cognitive-science based category theories used to formally describe re-

[1]Having acknowledged the distinctions above, we use *avatar* going forth.

[2]We follow role-playing conventions here, but recognize the distinction between gender and sex. There are multiple models of gender going far beyond male and female gender binary – this is an important area for future work.

sults from our models. Section III details *AIRvatar* as a system, describes the avatar creator that we developed, and outlines previous work accomplished with it. Section IV describes the data collection and user-study procedure. Section V details the methods and experimental design. Section VI contains the results and analysis of our experiments. We discuss the implications of our findings in Section VII, and describe limitations and proposed future work in Section VIII. Finally, we conclude with closing remarks in Section IX.

II. BACKGROUND

A. Avatars, Virtual Identities, and Player Identities

To formally describe the interrelationship between virtual and real-world identities, we begin with James Gee's definition of a third type of identity, which he terms the "projective identity" [7]. Gee describes the projective identity as a manifestation of values associated with both the player and the avatar. In this paper, we use Harrell's notion of a "blended identity" [8], which describes digital self-representations as *selective projections* of some aspects of a real player (e.g., preferences, control, appearance, personality, understanding of social categories, etc.) onto the actual implemented virtual representation. This includes the computational data structures that are used to implement and create them. Hence, we believe that studying these underlying data structures can consequentially be used to reveal aspects of a player's real-world identity.

B. Player Modeling & Game Data Mining

The use of game telemetry data mining in research often involves discovering patterns in the data to help designers to gain insight into the behaviors exhibited by players within the game. The type of models used for our analysis and prediction can be formally described using the taxonomy of player modeling approaches, introduced by computer scientists Adam Smith et al. The taxonomy consists of the *domain* (game actions or human reactions), *purpose* (generative or descriptive), *scope* (individual, class, universal, or hypothetical), and *source* (induced, interpreted, analytic, or synthetic) [9]. In this case, our models can be described as being universal, descriptive models of game actions from both induced, interpreted, and synthetic sources. Previous research in computational intelligence has shown how players' gameplay behavior reveals information about their real-world identities and behavior [4], personalities [10], [11], and motivational traits [12], [13].

The goal is often to categorize the large collection of player data into smaller discrete categories in a process called *clustering*. Behavioral clustering enables computational models of players to be constructed, enabling designers to develop approaches to better support the players such as player-adaptive designs, quantitatively evaluating user performance, and improving player experience and satisfaction [14]–[16]. Videogame researchers Drachen et al. did a comparison of the various common approaches for clustering *World of Warcraft* players based on level progression [17]. Each approach produced different clusters, based on their interpretability, distinction from one another, whether they depicted legal/possible representations in the game, and how representative of the original data set they were. Our choice of archetypal analysis (AA) was based on seeking interpretable and distinct clusters to be able to reason about the types of behaviors, values, and preferences being exhibited by players with their avatars.

C. Archetypal Analysis

Archetypal analysis (AA), introduced by Cutler and Breiman [18], is a method for reducing the dimensionality of multivariate data [19]. Given a set of multivariate data points, the aim of AA is to be able to represent each data point as a *convex combination* of a set of key data points called **archetypes**. For example, applying AA on a dataset of basketball players and their statistics [20] computationally revealed and represented the following four archetypes – "benchwarmer," "rebounder," "three-point shooter," and "offensive player." Every individual player in the entire data set could then be represented as a hybrid mixture of these archetypes [21]. Formally, given a data set of points $\{x_1, x_2, ..., x_n\}$, AA seeks to find a set of archetypes $\{z_1, z_2, ..., z_k\}$, where:

$$z_j = \sum_{i=1}^{n} \beta_{ij} x_i \qquad (1) \qquad\qquad \hat{x}_i = \sum_{j=1}^{k} \alpha_{ji} z_j \qquad (2)$$

Equation 1 means each archetype z_j resembles (i.e., represented using the same feature variables) as the data and Equation 2 specifies that each data point x_i can then be represented as a weighted combination of the archetypes. The objective function minimizes the residual sum of squares:

$$RSS = ||x_i - \hat{x}_i||^2 = ||x_i - \sum_{j=1}^{k} \alpha_{ji} z_j||^2 \qquad (3)$$

under the constraints that the weights $\sum \beta_{ij} = 1 \; \beta_{ij} \geq 0$ and coefficients $\sum \alpha ji = 1 \; \alpha_{ji} \geq 0$. These ensure the archetypes *meaningfully resemble* and are *convex mixtures* of the data. These archetypes are located on the data convex hull [18] and are represented as combinations of individual points, making them more easily interpretable [19], unlike other dimensionality reduction techniques like principal component analysis [22] and non-negative matrix factorization [23]. AA has been shown to be effective compared to other techniques for various AI-related problems, such as game recommendation systems [24].

D. Cognitive Categorization & Sociology of Classification

We use categorization models from cognitive science in order to formally describe the phenomena obtained from our models. We provide a brief introduction and discuss several of these models. In cognitive science, modern theories for categorization and classification are based on identifying members that are deemed "better examples" of a category than others, which are termed *prototypes* by psychologist Eleanor Rosch [25]. Thus, categorization of individuals occurs based on their *perceived distances relative to these prototypes*, termed "centrality gradience," [26] by cognitive scientist George Lakoff. He introduces the notion of "prototype effects," based on the theory that categorization is a cognitively-grounded and imaginative process involving metaphorical projection [26]. We use the following definitions from [27] to help describe individuals modeled in our dataset using archetypal analysis. Prototypes are described as the "best example" members of categories, while "marginalization" is a result occurring where members of a marginal category exist outside of social groups, or are less prototypical members of communities, or characterized by an individual having *multiple memberships*.

III. AIRvatar

We created an avatar customization system in the setting of a retro-styled fantasy role-playing-game (RPG)[3] (See Figure 1) with *AIRvatar* as its backend. Before customizing their avatars, players were presented with an introductory videogame opening sequence, providing a traditional role-playing game setting to help contextualize the style of the assets and motivate players to create avatars to be used as part of a videogame, not just a stand-alone avatar construction application. Players had two choices of avatar genders to select from, in turn providing access to a gender-specific base image and assets across the five customization categories for the visual appearance of their avatars – hair, head, body, arms, and legs. In each category, several sub categories of assets gave players more fine-grained control over their avatar's appearance, e.g., gloves for hands or pads for shoulders under the arm category. Players were provided an animated preview of their avatars and could rotate the view in any of the four directions. Each created character is 32×48 in pixel dimensions with a four-framed walkcycle animation for the front, back, left, and right-facing directions.

A. Previous Work with AIRvatar

We provide a brief overview of our previous research work conducted with *AIRvatar*. We studied what these constructed avatars revealed about players' values and perception based on the different technical components [28] of identity representation systems. Using Non-Negative Matrix Factorization (NMF), we analyzed the textual descriptions, visual images, and numerical statistics allocated to avatars by players [29]. It revealed how players used text differently for describing their avatars (e.g, for describing family relations, location information, or individual characteristics). We found that the allocation of statistical attributes by players could be computationally modeled using archetypal analysis to resemble well-known archetype roles of computer role-playing games (RPG) such as physically-dominant fighters, intelligence-oriented mages, and charm-oriented thieves [30]. While the focus on previous work has been on using the virtual identity (constructed avatar) to reveal aspects of the players' identity and values, the focus of this paper is different in that we are modeling the behaviors exhibited by players from telemetry data collected from using the system. Preliminary work in studying the data of a smaller set of players revealed that players exhibited different behaviors while customizing their avatars based on their gender [31]. Hence, in this paper, we seek to (1) computationally model these behaviors by players, (2) use these computational models to quantitatively assess how these behaviors differ, and (3) validate the model in predicting players' real-world genders, which are aspects of their real-world identities.

IV. Data Collection

We conducted a user study with participants from the social news and discussion site *Reddit* (`/r/samplesize` sub-reddit) who created characters using our avatar creator. Participants signed a consent form approved by the human subjects research committee at MIT. Participants were informed about the research aims of the project and that anonymous analytical data would be collected during avatar customization.

[3] Art assets from the *Mack Looseleaf Avatar Creator* (geocities.jp/kurororo4/looseleaf/) and the *Liberated Pixel Cup* (lpc.opengameart.org/)

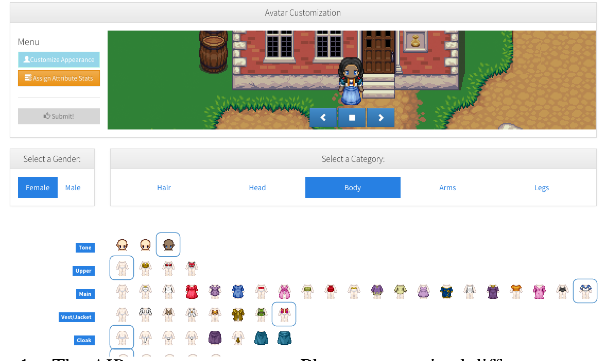


Fig. 1. The *AIRvatar* avatar creator. Players customized different components of their avatars, including multiple accessories and color variants based on the chosen avatar gender. An animated preview of the avatar in a walk-cycle could be toggled, including rotation options for four different perspectives.

V. Methods

A. Data collection and analysis of raw results

We first analyzed the data to gain a preliminary overview of the relationship between players' genders and behaviors.

1) Feature Variables: Our data set consisted of 12 feature variables, categorized into the following 4 categories:

- t_{Total} (total session duration)
- $t_{Category_i}$ (time spent customizing a item category) $Category_i \in \{Hair, Head, Body, Arm, Leg\}$
- t_{Gender_j} (time spent customizing with avatar gender) $Gender_j \in \{Female, Male\}$
- $t_{Orientation_k}$ (time spent in rotation orientation) $Orientation_k \in \{Front, Right, Left, Right\}$

2) Significant feature variables between player genders: To investigate if these feature variables differed between players based on gender, we performed multiple pairwise t-tests between the means of each feature variable for male and female players. Results were deemed significant at $p < .05$.

B. Constructing the behavioral archetypes model

We describe how archetypal analysis (AA) was used to develop our behavioral archetypes model.

1) Data preprocessing: We first converted all feature variable values into ratios. For feature variables 2–12, we divided their values by their respective t_{Total} values. We normalized t_{Total} by dividing it by the maximum value observed.

2) Determining number of archetypes: In order to determine the optimal number of archetypes, we varied the number of archetypes to be between $1 \leq k \leq 10$ as per [20]. AA was repeated 5 times for each value of k. We then used the residual sum-of-squares (RSS) to guide the selection of the optimal value of k, while considering the possibility of overfitting.

3) Selecting an optimal model: To balance the trade-off between model complexity and its RSS performance, we selected an optimal model by plotting the RSS errors against each value of k and used the Cattell scree test [32] to identify values of k along the "elbow" of the plot. We selected the smallest k that distributed (1) the number of players across archetypes and (2) the genders of players within clusters as similarly to $k + 1$ (i.e., the smallest changes to player distributions.)

C. Evaluating the behavioral archetypes model

We describe the procedure used to evaluate what each resultant behavioral archetype from our models represented.

1) Interpreting behavioral archetypes from β-weights: The β-weights enable us to decompose each behavioral archetype as a convex combination of the 12 original feature variables. We can then meaningfully interpret each behavioral archetype.

2) Identify clusters from α-coefficients: The α-coefficients defines the centrality gradience (defined in Section II-D) of each individual with respect to each behavioral archetype. We categorize players to archetypes with the highest α-coefficient.

3) Analyzing the distribution player and avatar genders across clusters: Having an understanding of each behavioral archetype (step 1) and being able to categorize individuals (step 2), we analyze how players are distributed across the behavioral archetypes (player-avatar demographics) according to (a) player gender and (b) avatar gender. We also consider the player-avatar demographics for marginalized individuals.

D. Model Validation by Predicting Player Gender

We validated our model by performing the task of predicting players' gender using only behavioral data. It would validate our hypothesis that players' behaviors during customization could reveal aspects of their real-world identities.

1) Supervised Learning using Support Vector Machines: To perform supervised learning, we made use of support vector machines (SVM) to train and test our model over the data set of players. Players were represented by a feature vector of the α-coefficients obtained from the AA model. We used a grid search to determine optimal values for the tunable parameters of the SVM – (1) the Gaussian kernel parameter γ and (2) cost C via grid search. The range of values for both were in the range $[10^{-5}, 10^{-4}, \ldots 10^{4}, 10^{5}]$. The performance of the best model selected was evaluated using 10-fold cross validation.

2) Model comparisons: We compared our model using two sets of experiments. The first set of experiments compared the behavioral archetype model against 4 control models – three of which corresponded to the groups of data (e.g., item category, gender, rotation orientation), and a fourth brute-force model with all 12 feature variables. The second set of experiments compared the performance of the behavioral archetype model with different values of k. This was to investigate the impact of the number of archetypes on model prediction performance.

VI. Results & Analysis

A. Player Demographics

Out of the 190 participants – 104 participants (54%) identified as "Male", 80 (43%) identified as "Female", and 6 (3%) listed "Other." This gave a fairly representative distribution between genders. For age-groups, 154 participants (80%) were between "18-24" years old, 32 (17%) were between "25-34" years old, and the other age groups were $< 1\%$.

B. Player Behavior Descriptive Statistics

Table I shows the descriptive statistics of the 12 feature variables that were tracked for all players using *AIRvatar*. We omitted 1 player with unusually long time durations, possibly from leaving the system running idle while away from the computer for a final data set of size $N = 190$. We can see that players spent about 10 minutes on average customizing their avatars. From the table, while we see that players customizing

TABLE I. Table showing the time duration spent by players in different aspects of our RPG customization system.

	Feature	min (s)	max (s)	mean (s)	sd (s)
Total	1. t_{Total}	150	5600	610	570
Item Category	2. t_{Hair}	0	470	95	85
	3. t_{Head}	0	2620	130	230
	4. t_{Body}	7.5	2290	190	21
	5. t_{Arm}	0	530	44	54
	6. t_{Leg}	0	4000	130	290
Avatar Gender	7. t_{Female}	0	504	340	380
	8. t_{Male}	0	290	250	590
Animation Orientation	9. t_{Front}	70	5500	570	540
	10. t_{Right}	0	290	17	38
	11. t_{Back}	0	160	23	37
	12. t_{Left}	0	190	8	18

TABLE II. Table showing the time duration spent by players in different aspects of our RPG customization system.

	Feature	mean (s) (Female)	mean (s) (Male)	p-value
Total	1. t_{Total}	735	530	$< .05$
Item Category	2. t_{Hair}	124	74	$< .05$
	3. t_{Head}	181	96	$< .05$
Avatar Gender	7. t_{Female}	678	96	$< .05$
	8. t_{Male}	400	415	$< .05$
Animation Orientation	9. t_{Front}	683	484	$< .05$

`Female` avatars spent more time on average than players customizing `Male` avatars, a Welch's t-test revealed that this difference was not statistically significant. Players spent the most amount of time customizing the avatar `body`, followed by `leg` and `head` customization joint-second, `hair`, and finally the `arm`. These differed compared to findings by Ducheneaut et al. who found that players in *Maple Story* and *World of Warcraft* (WoW) rated "hair-style" as most important. However, our findings were similar to Ducheneaut's findings players in *Second Life* who had high importance for "legs/torso". This was surprising because the style and setting of the assets in our system were closest to *Maple Story* and WoW, while furthest from *Second life*. Yet, our players' behaviors matched those of *Second Life*. We hypothesize that the demographics of the players, along visual style, factor into exhibited behaviors within a virtual system like ours.

C. Differences in Player Behavior based on Player Gender

We performed Welch's t-tests for each of the above features with player gender as the independent variable. Table II shows the descriptive statistics of the 6 features that had significant differences ($p < .05$) between their means. Female players (M= 735, SE= 81) spent a longer time customizing than male players (M= 530, SE= 40). This phenomena is consistent for all significant features except for t_{Male}, where male players spent longer than female players. One notable result is the different player behaviors exhibited when customizing characters of opposite genders – i.e., male players (M= 96, SE= 22) spent significantly less time than female players (M= 400, SE= 18) when customizing avatars of the opposite gender.

D. Implications of Number of Archetypes on Representation

Based on our model selection criteria described in Section V-B, we found that the optimal number of archetypes to be between $3 \leq k \leq 5$. We discovered that (1) for $k \geq 3$, a large majority (70%) of female players were always found on a single archetype and (2) as k increases, male players get further distributed across the remaining archetypes. This implied that

TABLE III. THE CHART DESCRIBES HOW EACH BEHAVIORAL ARCHETYPE WAS INTERPRETED BASED ON THE β-WEIGHTS.

	Name & Interpretation	male	female	hair	head	body	arm	leg	front	right	back	left	time	F	M	O	T
A1	**Male-preferring Non-styler**: Prefers `male` avatars, focused on `leg`, and doesn't rotate character. Mainly consist of "male" players.	M	L	L	L	L	L	*H	*H	L	L	L	L	6	37	2	45
A2	**Female-preferring Hair/Body Styler**: Strongly prefers `female` avatars, high focus on `hair` and `body`, and with a lot of time spent inspecting the `back` view. Mainly consist of "female" players.	L	H	H	L	H	M	M	M	L	H	M	M	70	19	2	91
A3	**General Styler**: No avatar gender preference. High focus on `hair` and rotating the avatar. Consist of "Male" players only.	M	M	*H	L	L	M	M	L	*H	*H	*H	M	0	7	0	7
A4	**Female-preferring Face Styler**: Strongly prefers `female` avatars. High focus on `head`. Highly engaged. Focused on side-profile rotation. Relatively equal split between "male" and "female" players.	L	*H	L	*H	L	L	L	M	H	M	M	*H	3	2	0	5
A5	**Male Physique Styler**: Strongly prefers `male` avatars. High focus on `body` & `arm`. Highly engaged. Focused on side-profile rotation. Mainly "male" players. More players self-identified as "other" than "female".	*H	L	L	M	*H	*H	L	M	H	M	M	H	1	39	2	42

Key: {T: Total, F: Female, M: Male, O: Other, L: Low, M: Medium, H: High, *: Indiciates highest β-weight value for given category.}

female players exhibited more consistent prototypical behaviors while male players exhibited more varied behaviors, with each behavior represented and characterized by a behavioral archetype. These results are shown in Table IV. Hence, the implication of varying k is to balance the trade-off between the expressibility (i.e., more archetypes $\rightarrow$ more expressibility) and the complexity (i.e., more archetypes $\rightarrow$ more difficult to interpret) of our models. To illustrate this, we visualize the data set for $k = [3, 4]$ archetypes as shown in Figure 2. In Figure 2(a) with $k = 3$ archetypes, we observe that male and female players are each closely associated with Archetype 2 and Archetype 3 respectively. Archetype 1 appears to have only male individuals associated with it. In Figure 2(b) with $k = 4$ archetypes, we now see that the few female players, originally in the center of, are now distributed across formed Archetype 4. After experimenting with various values, we decided on $k = 5$ as the optimal value for our purposes.

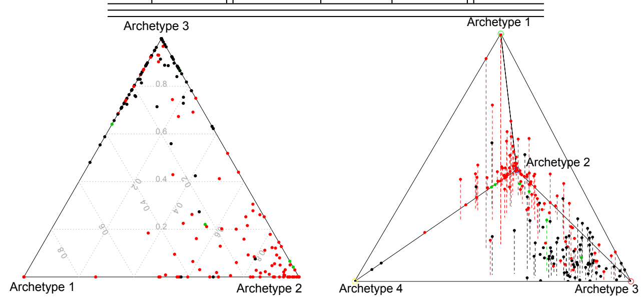
E. Interpreting Behavioral Archetypes

Having decided on $k = 5$ archetypes, our next step was to interpret each archetype using the α-coefficients, which define each archetype in terms of the original feature variables. We discretized the range of values of of each α-coefficient into low (L), medium (M), and high (H) and constructed Table III to aid with the interpretation. Behavioral archetypes differed in three aspects: (1) the preference for either male ("male-preferring") or female avatars ("female-preferring"), (2) the focus on different customizable parts of the avatar (e.g., "hair-styler", "body-styler", etc.,) and (3) the amount of time spent rotating the avatar's viewing profile. We also discuss the distribution of players within each archetype cluster based on players who used avatars with the same gender ("gender-synchronizing") or different gender ("gender-bending") as their own real-world genders. The total time by players within the system indicated how "engaged" the players were.

We found two different behavioral archetypes for players who *strongly preferred female avatars*. From Table III, we observed that A2 had a high percentage of female players and a strong preference on their avatars' `hair` and `body` ("hair-body styler"). A4 was also "female avatar-preferring," but was more evenly split between player genders. It exhibited strong preference for customizing the avatar's `head` ("face styler")

TABLE IV. TABLE SHOWING DISTRIBUTION OF PLAYERS ACROSS ARCHETYPES ACCORDING TO THEIR GENDER.

		Player Gender			
k	Arch.	Female	Male	Other	Total
2	A1	62	37	2	101
	A2	18	67	4	89
3	A1	0	4	0	4
	A2	6	80	4	90
	A3	74	20	2	96
4	A1	0	4	0	4
	A2	4	2	0	6
	A3	69	23	2	94
	A4	7	75	4	86
5	A1	6	37	2	45
	A2	70	19	2	91
	A3	0	7	0	7
	A4	3	2	0	5
	A5	1	39	2	42
6	A1	0	5	0	5
	A2	2	33	1	36
	A3	0	6	0	6
	A4	6	36	2	44
	A5	69	22	3	94
	A6	3	2	0	5



(a) $k = 3$ Archetypes (b) $k = 4$ Archetypes

Fig. 2. The plots above visualize the individuals in the data set in terms of the archetypes using a barycentric coordinate system. Symbols indicate player gender where: ● female ● male ● other. When $k = 3$ in (a), we observe that female players load more on A3 and male players on A1. But a number of male players also load highly on A3. By increasing to $k = 4$ in (b), the additional archetype is formed from female players originally in the center.

and high engagement from time spent during customization. Both archetypes were characterized by high time spent viewing their avatars from different angles, with high emphasis on viewing from the `back` (A2) or the `right` (A4) of the avatar.

The behavioral archetype of *strongly preferring male avatars* (A5) appeared to be skewed heavily toward male players only ("male-preferring"). The archetypal behavior for these players were being highly engaged and exhibited "physique styler" traits (focus on `body` and `arm` and to a lesser extent the `head`.) In comparison, the behavioral archetype (A1), which had was slightly "male-preferring", had high focus on the lower-body region (e.g., `leg`) but did not appear to spend any time rotating their characters at all. Finally, the behavioral archetype A3 showed *no avatar gender preference* by players, but had highly "male gender-synchronized" players that were "hair stylers". These players spent large amounts of time viewing their avatars from different rotation orientations.

F. Demographic Patterns with Archetypal Clustering

With our understanding of each of the $k = 5$ behavioral archetypes, we take a closer look at the demographics of the clusters of players both near and far away from each archetype. This is shown in Table V. A random sample of 10 marginalized (i.e., located furthest from all archetypes) individuals are in row `M`. Our discussion focuses on our observation that clusters around archetypes were (1) either player gender-diverse (A1, A2) or player gender-skewed (A3, A4, A5) and (2) either male avatar (A1, A3, A5) or female avatar (A2, A4)-oriented.

We observed that the cluster around A2 represented a large gender-diverse cluster of players that preferred female avatars. It had the biggest cluster of players located closest ($\alpha_k \geq .8$) to it, and all were creating female avatars. The gender split was fairly diverse between male (36%) and female (64%) players. Also worth noting is that A2 had no pure archetypes ($\alpha_k = 1$). In contrast, clusters around A1 are player gender-diverse but very small in size, and consists of two pure archetypes. Both these clusters appear have high "gender-bending" players who use avatars not matching their real world genders.

We observed that clusters around A3 and A5 were skewed toward "male avatar-preferring" male players, while clusters around A4 are heavily skewed toward "female avatar-preferring" female players . Each of these archetypal clusters also possess a single pure archetype. Hence, in comparison to clusters A1 and A2, these clusters represent players who prefer playing with avatars matching their own gender ("gender-synchronizing" behaviors). This highlights the importance of the interpretation steps in Section V-C to evaluate behavioral archetypes based on both player and avatar genders. Both A1 and A3 clusters were male avatar-oriented with 3 top individuals, but players in A1 "gender-bended" while those in A3 were "gender-synchronized." The marginalized cluster (N=30) possessed an even distribution of male and female players (50%), with a slight skew toward female avatars (63%). Most players had high α-coefficients for A1 and A2.

G. Predicting Player Gender from Behavioral Archetypes

Our next step was to validate our models with a practical task, informed from the results in previous sections. We felt that these behavioral models would be effective in predicting players' gender from just in-game behavioral data. Our model achieved an accuracy score of F_1= .81 with a relative absolute error of e_{rel}= .34. The resulting confusion matrix from this model is shown in Table VI. The model had a higher precision rate for predicting `male` (.88) versus `female` (.76) players, but a higher recall rate for predicting `female` (.91) versus `male` (.80) players. These point toward the effects of an uneven distribution of player genders in our data set, highlighted more as `other` was never predicted. To further

TABLE V. TABLE WITH TOP INDIVIDUALS ($\alpha \geq .80$) OF ARCHETYPES (A1-A5) AND A RANDOM SAMPLE OF MARGINALIZED INDIVIDUALS (M).

	α_1	α_2	α_3	α_4	α_5	Player Gender	Avatar Gender
A1	**1.00**	0.00	0.00	0.00	0.00	Female	male
	1.00	0.00	0.00	0.00	0.00	Male	male
	0.80	0.03	0.01	0.06	0.10	Male	male
A2	0.00	**0.98**	0.02	0.00	0.00	Female	female
	0.00	**0.92**	0.04	0.02	0.02	Female	female
	0.00	**0.90**	0.07	0.00	0.03	Male	female
	0.03	**0.90**	0.01	0.06	0.00	Male	female
	0.03	**0.87**	0.01	0.08	0.00	Female	female
	0.00	**0.86**	0.08	0.07	0.00	Female	female
	0.08	**0.85**	0.00	0.00	0.07	Female	female
	0.00	**0.85**	0.03	0.12	0.00	Female	female
	0.00	**0.84**	0.00	0.01	0.14	Female	female
	0.07	**0.83**	0.04	0.05	0.00	Male	female
	0.00	**0.83**	0.00	0.00	0.17	Female	female
	0.10	**0.82**	0.07	0.00	0.00	Male	female
	0.03	**0.81**	0.05	0.06	0.05	Male	female
	0.00	**0.81**	0.06	0.00	0.13	Female	female
A3	0.00	0.00	**1.00**	0.00	0.00	Male	male
	0.05	0.00	**0.95**	0.00	0.00	Male	male
	0.00	0.00	**0.89**	0.11	0.00	Male	male
A4	0.00	0.00	0.00	**1.00**	0.00	Female	female
	0.15	0.00	0.00	**0.85**	0.00	Female	female
A5	0.00	0.00	0.00	0.00	**1.00**	Male	male
	0.08	0.00	0.00	0.01	**0.91**	Male	male
	0.00	0.01	0.00	0.09	**0.90**	Male	male
	0.10	0.00	0.00	0.01	**0.89**	Other	male
	0.14	0.00	0.00	0.00	**0.86**	Male	male
M	**0.42**	**0.57**	0.00	0.01	0.00	Male	female
	0.26	**0.20**	**0.31**	0.07	0.17	Male	male
	0.27	**0.26**	0.03	0.08	0.36	Male	female
	0.30	**0.62**	0.00	0.08	0.01	Male	female
	0.31	**0.59**	0.00	0.10	0.00	Female	female
	0.42	**0.45**	0.00	0.10	0.03	Female	female
	0.00	**0.66**	0.00	**0.34**	0.00	Female	female
	0.04	**0.44**	**0.15**	**0.38**	0.00	Female	female
	0.44	**0.26**	0.02	0.01	**0.27**	Male	male
	0.69	0.00	**0.28**	0.03	0.00	Male	male

TABLE VI. THE CONFUSION MATRIX RESULTING FROM SUPERVISED LEARNING USING OUR BEHAVIORAL ARCHETYPES MODEL. PREDICTION ACCURACY FOR "FEMALES" WAS HIGHEST, FOLLOWED BY "MALES". "OTHER" WAS NOT PREDICTED, OFTEN BEING MISTAKEN FOR "MALE".

	Predicted		
	Female	Male	Other
Female	73	7	0
Male	21	83	0
Other	2	4	0

evaluate the implications of these results, we performed two set of experiments – (1) comparison between our behavioral archetype model against the four control models and (2) comparison between different behavioral archetype models based on different values of specified archetypes k.

1) Archetype Model vs. Control Models: Based on accuracy measures, we observed that the AA model (F_1= .82) did significantly better than both Control #3 (F_1= .52) and Control #4 (F_1= .58). Our model only under-performed compared to both Control #1 (F_1= .83) and Control #4 (F_1= .83). These results highlight the importance of the time spent customizing as either female or male avatars in revealing players' real-world genders. It also shows that the AA model is robust in retaining the explanatory power of the original feature variables.

2) Comparing Archetype Models: The model performed significantly worst ($F_1 = 0.69$) for $k = 2$. For $k > 2$, the best performance was from the model with $k = 3$ ($F_1 = .83$), followed by our chosen model with $k = 5$ ($F_1 = .81$). This backs our choice to use a higher number ($k = 5$) of archetypes as it only resulted in a small trade off in classification accuracy, but enabled us to have a higher expressibility for interpreting archetypes. Increasing k reduces accuracy and increase the complexity (i.e., more difficult to interpret) of archetypes.

VII. Discussion

From our results, we discovered that players' real-world gender can be computationally modeled using their virtual behaviors, specifically, during the avatar customization process. Here we discuss implications of these findings. We consider the importance of avatar customization systems being able to adequately support the nuances exhibited by different players based on their real-world identities. While our focus here is on player gender, we can extend these notions to other aspects of players' identities that are similarly affected by their avatars.

A. The influence of gender on time spent customizing avatars

Our results highlight that the time players spent customizing their avatars was the most importance factor in revealing an aspect of the players' identity (e.g., gender). They revealed that not only was there a difference between male and female players in terms of the time spent customizing avatars, but also that there was a nuanced difference depending on whether players were customizing an avatar with the same or different gender to themselves. For example, we observed that both male and female players spent on average of six minutes when customizing male avatars, but female players spent seven times longer than male players when customizing female avatars. The implications of these findings point toward the need to consider the impact of how providing or restricting the aspects of avatar representation may affect players' attachment to their avatars, and subsequently, their engagement within the system or the application for which the avatar would be deployed.

B. Gender distributions and archetypal behavioral patterns

From the five behavioral archetypes, we observed that two archetypes had an emphasis on female avatars (A2, A4), two archetypes had an emphasis of male avatars (A1,A5), while the fifth had no discernible emphasis between either avatar gender (A3). This implies that avatar gender distribution was fairly even between archetypes. For player gender, however, the distribution was uneven. Even with an increasing number k of archetypes, the same proportion of female players was always associated with one archetype, while male players become more distributed across the remaining archetypes. The suggests that female players might demonstrate more consistent behavior when customizing avatars than male players who exhibited more varying behaviors. For example, from archetypes A1, A3, and A5, we see that male players differ in the following archetypal behaviors: (1) spending little time in general and only focusing on the `leg`, (2) having no avatar gender preference, focusing on the `hair` and `body` and considering their avatars' appearance from different rotation angles, and (3) having a strong preference for male avatars, with an emphasis on the `body` and `arm` while being highly engaged in the system. Female players, however, generally exhibited consistent behavior according to A2 – with a high preference for `Female` avatars, emphasis on the avatar's `hair`, and viewing their avatar from the `back`-view rotation.

C. Fidelity and importance of customization components

Our results highlighted how players valued the components (e.g., `head`, `hair`, `body`, `arm` and `leg`) of the avatar differently based on the gender of the avatar. It also revealed the implicit relationship between pairs of components (1) `hair` and `body` and (2) `body` and `arm` (male avatars). A reason may be due to the different number of customization sub categories and assets available to each avatar gender. Also, the two categories that ranked highly on > 1 archetype were `hair` (A2, A3) and `body`. These results bear similarities to related research that showed that players ranked hair style and color as the most important feature, which players spent the most amount of time customizing. Ducheneaut et al. hypothesize that this is may be due to two reasons: (1) due to it being a "malleable" part of a real human body that can besed to build one's appearance and identity and (2) due to its "visibility," compared to other parts often covered with accessories.

VIII. Limitations & Future Work

A. Target demographics of participants

The participants for this study came from the social news site Reddit (`/r/samplesize`), which inevitably skewed the data set toward younger, more technology-savvy individuals. This was appropriate for our aims to gain insight into the everyday behaviors of people, which did not necessarily require them to play a lot of games. A broader study distinguishing between self-identified gamers and non-gamers would be worth studying, as more gaming oriented players might have different reasons to customize their avatars (e.g., an inclination to explore a different gender). An upside was that we obtained a fairly diverse population of users gender-wise. In the real world, gender is more complex than in this avatar creation system. Yet, people negotiate systems with these limitations all the time. In the future, a more gender-diverse demographic of users (or users aware of a greater range of gender diversity) could be used as normative categories to investigate the impacts of imposing fewer gender norms in the system.

B. Genre, style, and fidelity of game and assets

Another aspect that might have influenced the behaviors of players is the genre and style of the avatar customization interface. The setting was a traditional RPG setting, with a more fantasy-oriented 16-bit art style (that would now be considered "retro" graphics). We suspect that given a different genre or setting (e.g., space exploration, social simulation) of videogames, the behaviors of these players would differ. We are currently working to integrated *AIRvatar* into a different avatar customization interface based on a more realistic setting. Also, the assets used differed in the number of options between avatar genders, and some assets provided different color variations, while others only had one. Future work could balance out the available options to control on system fidelity.

C. Experimental Design and Data Collection

In a future study, asking players to construct avatars of both genders could provide greater insight into the behavioral differences that players might have when creating avatars with differing gender identities. The upside to our current approach was that we could concretely model player behaviors that demonstrated stronger inclinations toward a particular avatar gender. A within-subjects study might yield insight into the variations in a player's behavior when customizing avatars of different identities. We also collected information on alternate aspects of player identities such as their personality profiles. Preliminary results using the BIG-5 International Personality Item Pool (IPIP) showed several significant, but small correlations with the behavioral data. Our archetypal behavior model showed promising results in relation to player "conscientiousness." Finally, in our current experimental design, participants could not use their constructed avatars in an actual game setting. Recent work has demonstrated the effect of virtual identities on learning in games [33]. Thus, we intend to study the effectiveness of behavioral archetype models in assessing performance within in-game environments and tasks.

IX. Conclusion

In this paper, we have described an approach to computationally model players' identities using behavioral data collected during the avatar customization process. We used archetypal analysis to develop these models, which enabled us to identify several prototypical behavior patterns exhibited by players such as (1) the preference, or lack thereof, for particular avatar genders, (2) the different relative levels of importance assigned to different customizable components of avatars, and (3) the impact of both player gender and avatar gender in affecting the aforementioned behavioral patterns. Additionally, we demonstrated how our behavioral archetype model had high accuracy ($F_1 = 0.81$) in predicting the player's real-world gender from telemetry data collected from within our system. Our results demonstrate the effectiveness of archetypal analysis for modeling several interesting phenomena, such as female players having more consistent behavioral patterns compared to male players, who demonstrated more variation in their usage patterns. These highlight the importance of considering how real-world and virtual identities affect one another, and how AI and computational intelligence approaches can be used to effectively model these behaviors for critical evaluation.

Acknowledgment

We thank the anonymous reviewers for valuable feedback. This material is based upon work supported by the National Science Foundation under Grant No. 1064495. We also acknowledge the support of a QCRI-CSAIL Collaboration Grant.

References

[1] Z. Waggoner, *My avatar, my self: Identity in video role-playing games*. McFarland, 2009.

[2] N. Yee, N. Ducheneaut, M. Yao, and L. Nelson, "Do men heal more when in drag?: Conflicting identity cues between user and avatar," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011, pp. 773–776. [Online]. Available: http://doi.acm.org/10.1145/1978942.1979054

[3] D. F. Harrell and S. V. Harrell, "Imagination, computation, and self-expression: Situated character and avatar mediated identity," *Leonardo Electronic Almanac*, vol. 17, no. 2, 2012.

[4] N. Yee and J. Bailenson, "The proteus effect: The effect of transformed self-representation on behavior," *Human communication research*, vol. 33, no. 3, pp. 271–290, 2007.

[5] N. Yee, "Motivations for play in online games," *CyberPsychology & behavior*, vol. 9, no. 6, pp. 772–775, 2006.

[6] N. Yee, N. Ducheneaut, H.-T. Shiao, and L. Nelson, "Through the azerothian looking glass: Mapping in-game preferences to real world demographics," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12, 2012, pp. 2811–2814.

[7] J. P. Gee, "What video games have to teach us about learning and literacy," *Computers in Entertainment (CIE)*, 2003.

[8] D. F. Harrell, *Phantasmal Media: An Approach to Imagination, Computation, and Expression*. MIT Press, 2013.

[9] A. M. Smith, C. Lewis, K. Hullet, and A. Sullivan, "An inclusive view of player modeling," *Proceedings of the 6th International Conference on Foundations of Digital Games*, pp. 301–303, 2011.

[10] S. Tekofsky, J. Van Den Herik, P. Spronck, and A. Plaat, "Psyops: Personality assessment through gaming behavior," in *Proceedings of the Foundations of Digital Games Conference*, 2013, pp. 12–65.

[11] N. Yee, N. Ducheneaut, L. Nelson, and P. Likarish, "Introverted Elves & Conscientious Gnomes : The Expression of Personality in World of Warcraft," *In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, pp. 753–762, 2011.

[12] A. Canossa, J. B. Martinez, and J. Togelius, "Give Me a Reason to Dig Minecraft and Psychology of Motivation," *In Proceedings of IEEE Conference on Computational Intelligence and Games*, 2013.

[13] N. Yee, "Motivations for play in online games." *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, vol. 9, no. 6, 2006.

[14] A. Drachen, R. Sifa, C. Bauckhage, and C. Thurau, "Guns, swords and data: Clustering of player behavior in computer games in the wild," in *Computational Intelligence and Games (CIG), 2012 IEEE Conference on*. IEEE, 2012, pp. 163–170.

[15] G. N. Yannakakis and J. Hallam, "Real-time game adaptation for optimizing player satisfaction," *Computational Intelligence and AI in Games, IEEE Transactions on*, vol. 1, no. 2, pp. 121–133, 2009.

[16] G. N. Yannakakis and J. Togelius, "Experience-driven procedural content generation," *Affective Computing, IEEE Transactions on*, vol. 2, no. 3, pp. 147–161, 2011.

[17] A. Drachen, C. Thurau, R. Sifa, and C. Bauckhage, "A comparison of methods for player clustering via behavioral telemetry," *arXiv preprint arXiv:1407.3950*, 2014.

[18] A. Cutler and L. Breiman, "Archetypal analysis," *Technometrics*, vol. 36, no. 4, pp. 338–347, 1994.

[19] C. Bauckhage and C. Thurau, "Making archetypal analysis practical," in *Pattern Recognition*. Springer, 2009, pp. 272–281.

[20] M. J. Eugster, "Archetypal athletes," *arXiv preprint arXiv:1110.1972*, 2011.

[21] S. Seth and M. J. A. Eugster, "Probabilistic archetypal analysis," arXiv.org, Tech. Rep., 2014. [Online]. Available: http://arxiv.org/abs/1312.7604

[22] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.

[23] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[24] R. Sifa, C. Bauckhage, and A. Drachen, "Archetypal game recommender systems," *Proc. KDML-LWA*, 2014.

[25] E. Rosch, "Principles of categorization," *Concepts: Core readings*, pp. 189–206, 1999.

[26] G. Lakoff, *Women, Fire, and Dangerous Things*. Cambridge University Press, 1990.

[27] G. C. Bowker and S. L. Star, *Sorting Things Out: Classification and its Consequences*. MIT Press, 1999.

[28] D. F. Harrell, "Computational and cognitive infrastructures of stigma: Empowering identity in social computing and gaming," *Proceedings of the 7th ACM Conference on Cognition and Creativity*, pp. 49–58, 2009.

[29] C.-U. Lim and D. F. Harrell, "Developing computational models of players' identities and values from videogame avatars," in *Proceedings of the 10th International Conference on the Foundations of Digital Games*, 2015.

[30] ——, "Revealing social identity phenomena in videogames with archetypal analysis," in *Proceedings of the 6th International AISB Symposium on AI and Games*, 2015.

[31] ——, "Toward telemetry-driven analytics for understanding players and their avatars in videogames," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '15. New York, NY, USA: ACM, 2015, pp. 1175–1180. [Online]. Available: http://doi.acm.org/10.1145/2702613.2732783

[32] R. B. Cattell, "The scree test for the number of factors," *Multivariate behavioral research*, vol. 1, no. 2, pp. 245–276, 1966.

[33] D. Kao and D. F. Harrell, "Toward evaluating the impacts of virtual identities on stem learning," in *Proceedings of the 10th International Conference on the Foundations of Digital Games*, 2015.