Toward Telemetry-driven Analytics for Understanding Players and their Avatars in Videogames

Chong-U Lim
MIT
77 Massachusetts Avenue
Cambridge, MA 02139 USA
culim@mit.edu

D. Fox Harrell
MIT
77 Massachusetts Avenue
Cambridge, MA 02139 USA
fox.harrell@mit.edu

CHI'15 Extended Abstracts, April 18–23, 2015, Seoul, Republic of Korea.
ACM 978-1-4503-3146-3/15/04.
http://dx.doi.org/10.1145/2702613.2732783

Abstract
A great deal of research has gone into understanding the relationships between social behaviors, player preferences, and identities in both real-world and virtual environments. In this works-in-progress report, we describe *AIRvatar*, a tool that telemetrically collects data such as session-based time durations and click events during the process of avatar customization within a videogame of our own creation. We present results from a user-study of 181 players, highlighting how social phenomena such as gender-related stereotypes of users can be revealed, particularly when players' self-identified real world genders contrast with the gender identities of their avatars.

Author Keywords
Computer Games; Telemetry; Avatars; Identity; Gender Representation

ACM Classification Keywords
H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities; I.2.1 [Applications and Expert Systems]: Games; K.4.2 [Social Issues]; K.8.0 [General]: Games.

Introduction
Digital media provide many ways for people to represent themselves [11], and this may occur explicitly through

customization choices (e.g., selecting the gender for a character in a virtual world) or implicitly through exhibited behavior and actions (e.g., spending more time customizing a character's hair color versus its shoe color.) Understanding users, their constructed avatars, and their behaviors in virtual environments is an active area of research in the HCI community [3, 17]. Typically, in this type of research two kinds of data sources are used: 1) self-reported data and 2) data mined in-game behavioral data (game telemetry). Self-reported data are used to evaluate subjective characteristics such as user experience, motivations, and personality profiles, through the use of standardized tests (e.g., BIG-5 IPIP, Reiss Motivation Profile), but have various limitations [7], are harder to evaluate and may be subject to survey bias [1]. Game telemetry data mining overcomes these shortcomings and adds explanatory power through providing direct access to large populations of user information, demographics, behaviors, and usage patterns [1, 2].

While previous work has focused on more limited types of identity phenomena and correlations, here we want to go beyond stating that "gender x demonstrates more of personality y using survey data." We want understand how nuanced types of social phenomena, such as stereotyping and privilege, can be implicitly revealed through player behavioral data collected telemetrically. Such social phenomena may reflect aspects of players' real-world identities and selves. Here, *real-world* is used to describe aspects of player's identity that exists outside of the game world and the distinction between real-world and virtual-world identities is of interest as people might create digital identities that are distinct from the ones they use in everyday life [15]. This research seeks to address the question of how telemetrically obtained data can be used to achieve such insight.

This works-in-progress presents an Artificial Intelligence (AI) system called ***AIRvatar***. Together with a system of our design, it combines telemetry and user analyses facilitated by the system to look at more nuanced types of phenomena. We shall demonstrate the effectiveness of the system by giving specific examples of the types of analyses we can produce, namely analyzing phenomena that occur when players' self-identified real world genders contrast with the gender identities of their avatars. As shown later, *AIRvatar* revealed gender-related stereotypes exhibited by female players when customizing their virtual avatars. Our goal is to work toward digital identity representation systems that adequately supporting users' needs and values regarding self-representation [10], to enhance learning, and to allow developers to more expressively model a diversity of identities. By considering both real-world and virtual identities, we seek to avoid designs that potentially reinforce undesirable social phenomena (e.g., discrimination, stereotyping) that may occur [9].

Related Work

Beyond Real vs. Virtual Identities

In [15], social scientist Nick Yee describes the "Proteus Effect," where he states that embodiment may lead to shifts in self-perception both online and offline based on the avatar's features or behaviors. To understand the relationship between player and avatar, James Gee's introduces the "projective identity" [6], which Harrell describes as "manifesting the ways that real player values are reconciled with values understood as being associated with avatars" [10]. In our work, we extend upon this to use Harrell's notion of a "blended identity." Under this view, most digital self-representations are 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. Hence, we consider multiple aspects of users' identities, including real-world selves and virtual representations, in our approach. Relating in-game behavior to real-world identities, such as demographic segments [16] and personality assessments [13], has demonstrated useful insight into understanding how to match game-based mechanisms to better support players. A common measurement scale for personality assessment is the Big-5 International Personality Item Pool [8], which measures Openness, Conscientiousness, Extraversion, Agreeableness, and Emotion Stability (OCEA-ES). Because these factors were derived from analyses of English words, there are concerns about the cross-cultural validity of these factors [5]. However, they have been shown to be reliable and moderately accurate for everyday life and online interactions [14], and are effective measures for understanding players in virtual worlds.

Revealing Gender-related Social Phenomena
In virtual environments, it is common for players to gender-bend (i.e., use avatars that are of an opposite gender). Yee demonstrated in [17] that often players' behaviors conform to stereotypes associated with their avatar's appearance (e.g., women preferring to heal,) but *only when using avatar characters of the opposite gender*. Other studies have studied the effects of the appearance of avatars, particularly in relation to overly sexualized appearances, which resulted in more body-objectifying thoughts in participants [4]. These are examples of "cross-stereotyping," in which players seem to attribute a more limited range of behaviors to other genders than they do to their own, identified as a common phenomenon in "identity tourism" [12]. We later demonstrate how such nuanced phenomena may be revealed with *AIRvatar*.

Approach

We hypothesize that *behavioral data collected telemetrically of players customizing virtual avatars in videogames can reveal social phenomena related to players' perception of genders both similar and different to their own.* We conducted a user study with participants from the social news and discussion site *Reddit*. A post was made on the `/r/samplesize` sub-reddit inviting users to take part in a videogame we created called *Heroes of Elibca*. It explained the research goals of the project and informed users that behavioral data would be collected anonymously for this purpose. The reason for using a videogame of our own creation was that, in the past, users had trouble creating characters if they did not know what they would be used for. We wanted a familiar videogame setting (i.e., computer role-playing game) to contextualize the style of the assets used in the customization system, though we did not provide functionality for using characters within a videogame.

Upon launch, players were presented with an introductory sequence to provide context for their player character's role in the game. Next, players were presented with an avatar customization interface[1] where they could customize the visual appearance of their characters (See Figure 1) across five main categories (hair, head, body, arms, and legs.) In each category, several sub-categories provided finer-grained levels of assets for customization. E.g., in the arms category, the shoulder, forearm, and wrists could have pads, gauntlets, and gloves. For the body, accessories such as swords and other physical features such as wings could be added. Several of these assets also had multiple color variants. Players also

[1]Art assets from the publicly available *Mack Looseleaf Avatar Creator* (`http://www.geocities.jp/kurororo4/looseleaf/`) and the *Liberated Pixel Cup* (`http://lpc.opengameart.org/`)

customize the statistical values of their player character across six commonly videogame attributes – strength, endurance, dexterity, intelligence, charisma, and wisdom. Each attribute could be assigned points on a 7-point Likert scale with a total of 27 points to be allocated. Once players had customized their characters, they completed the BIG-5 IPIP 100-item (5-point Likert scale) personality questionnaire and a demographic survey.

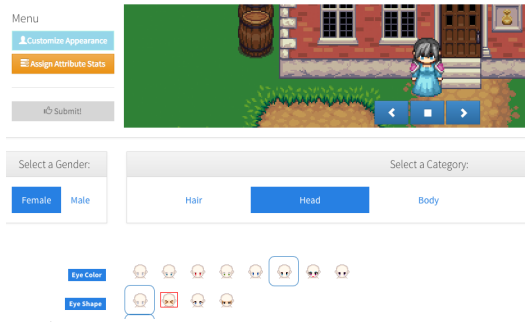

Figure 1: The customization interface of *Heroes of Elibca*.

Preliminary Results

In this section, we present preliminary results collected using *AIRvatar* from the user-study with *Heroes of Elibca*.

Demographic Information

For gender, 98 participants (54%) identified as "Male", 78 (43%) identified as "Female", and 5 (3%) listed "Other." This gave a fairly even distribution between genders for our sample size. For age-groups, 145 participants (80%) were between "18-24" years old, 31 (17%) were between "25-34" years old, and the other age groups were $< 1\%$.

Gender and Personality

We conducted one-way ANOVAs for each BIG-5 factor as dependent variables and Player Gender as the independent variable. We found significant differences between genders for "Emotional Stability" $(F[2, 178] = 5.72\ , p < .01)$ and "Agreeableness" $(F[2, 178] = 3.41\ , p < .05)$. Post-hoc Tukey's tests showed male players $(M = 68.5,\ SE = 1.46)$ had higher "Emotional Stability" scores than females $(M = 60.8,\ SE = 1.70)$, but males had lower "Agreeableness" scores $(M = 68.9,\ SE = 1.41)$ than female players $(M = 74.0,\ SE = 1.39)$.

Gender and Attribute Statistics

Figure 2 shows attribute point allocations made by female players for both male and female avatars. We observed that female players assigned significantly higher "Strength" and "Endurance" attribute points to male avatars than female avatars. They also assigned significantly higher "Intelligence" and "Wisdom" attribute points to female avatars than male avatars.

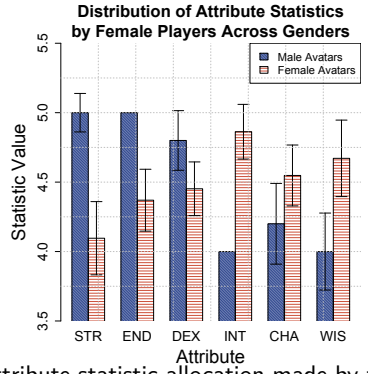

Figure 2: Attribute statistic allocation made by female players.

Timing and Click Events

We used Welch's t tests to compare the time durations and click event frequencies made players. We found significant effects of gender on time spent customizing "Hair" $(t(125) = 2.0, p < .05, Cohen's\ d = .30)$ and "Head" $(t(84) = 2.2, p < .05, Cohen's\ d = .34)$ categories of avatars. Female players spent more time customizing the "Hair" $(M = 120s,\ SE = 11s)$ and "Head" $(M = 180s,\ SE = 39s)$ categories than male players $(M = 71,\ SE = 6.4s)$ $(M = 91s,\ SE = 8.7s)$.

We omitted one player with unusually long time durations, possibly from leaving the system running idle while away. We found significant effects of gender on icon click events ($t(115) = 2.2,\ p < .05,\ Cohen's\ d = .33$) and color click events ($t(115) = 2.1,\ p < .05,\ Cohen's\ d = .32$) Females clicked more customization items ($M = 91, SE = 8.2$) and color options ($M = 71, SE = 8.7$) than males ($M = 70, SE = 5.2$) ($M = 51,\ SE = 4.4$).

Findings and Discussion

We discuss how using *AIRvatar* to telemetrically obtain data on fine-grain player actions and behaviors enabled us to measure and discover nuances in gender-related phenomena and preferences exhibited by players.

Social phenomena like gender-related stereotyping could be observed through players choices. Previous research has highlighted effects gender bending by players on their in-game behaviors [17] and our findings also demonstrated that gender-bending might give rise to gender stereotypes. Female players appeared to possess more stereotypical notions of masculinity and femininity. Many gave male avatars significantly more strength and endurance attribute points than female avatars, and significantly more intelligence and wisdom points for female avatars than male avatars. The combative-oriented of the RPG genre may be a contributing factor to this and we plan to use *AIRvatar* for alternate videogame genres and systems to explore this phenomenon further.

Gender-based personality differences could be observed through behavioral data. The personality assessments did not show significant difference for the "Conscientiousness" factor, but we demonstrated that *timing and click event analytics could be reveal such personality traits*. Female players were more engaged with the system than male players, as they spent almost twice as much time customizing the hair and heads of their avatars and explored more accessories and color options than their male counterparts. Such behaviors may be interpreted to reflect conscientiousness. This highlights how such behavioral data may provide more reliable assessments of personality than self-reported surveys, as related research has also shown [13].

Player preferences could be observed through behavioral data. We observed that players spent most time customizing the bodies, followed by head, hair, leg, and arms, of their avatars. This differs from other studies, where players often rated hair as the most important [3]. Hence, analyzing behavioral data provides insight not easily inferred from static media (e.g., final avatar image), highlighting the benefits of analytics-driven approaches for revealing implicit or hidden preferences of players.

Conclusions and Future Work

In this paper, we demonstrated the effectiveness of an analytics-driven approach for analyzing phenomena that occur in virtual environments, such as stereotyping that occurs from contrasting gender identities between players and their avatars. This extends other common methods, like correlating self-reported data, by obtaining data telemetrically, which is less obtrusive and enables insight to be obtained in a bottom-up manner through implicit behaviors exhibited by players. We seek to extend our *AIRvatar* to use AI-based techniques like clustering and natural language processing on other data collected like images and text input. Our goal is to enable players and designers of digital identity systems to be empowered for self-representations virtually in a manner that also prevents undesirable social phenomena from manifesting.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 1064495.

References

[1] 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. In *Proceedings of the IEEE Computational Intelligence and Games Conference*, IEEE (2012), 139–146.

[2] Drachen, A., Sifa, R., Thurau, C., and Bauckhage, C. A comparison of methods for player clustering via behavioral telemetry. In *Proceedings of 8th Annual Conference on Foundations of Digital Games* (2013).

[3] Ducheneaut, N., Wen, M.-H., Yee, N., and Wadley, G. Body and mind: A study of avatar personalization in three virtual worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2009), 1151–1160.

[4] Fox, J., Bailenson, J. N., and Tricase, L. The embodiment of sexualized virtual selves: The proteus effect and experiences of self-objectification via avatars. *Computers in Human Behavior 29*, 3 (2013), 930 – 938.

[5] Funder, D. C. *The personality puzzle.* WW Norton & Co, 1997.

[6] Gee, J. P. What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE) 1*, 1 (2003), 20–20.

[7] Goguen, J. A., and Linde, C. Techniques for requirements elicitation. *RE 93* (1993), 152–164.

[8] Gow, A. J., Whiteman, M. C., Pattie, A., and Deary, I. J. Goldberg's 'ipip' big-five factor markers: Internal consistency and concurrent validation in scotland. *Personality and Individual Differences 39*, 2 (2005), 317–329.

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

[10] Harrell, D. F. Toward a theory of critical computing. CTheory, ctheory.net/articles.aspx?id=641, 2010.

[11] Livingston, I. J., Gutwin, C., Mandryk, R. L., and Birk, M. How players value their characters in world of warcraft. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work* (2014), 1333–1343.

[12] Nakamura, L. *Digitizing Race: Visual Cultures of the Internet.* University of Minnesota Press, 2008.

[13] Tekofsky, S., Van Den Herik, J., Spronck, P., and Plaat, A. Psyops: Personality assessment through gaming behavior. In *Proceedings of the 8th International Conference on the Foundations of Digital Games.* (2013), 12–65.

[14] Yee, N. *The Proteus Paradox: How Online Games and Virtual Worlds Change Us-and how They Don't.* Yale University Press, 2014.

[15] Yee, N., and Bailenson, J. The proteus effect: The effect of transformed self-representation on behavior. *Human communication research 33*, 3 (2007), 271–290.

[16] Yee, N., Ducheneaut, N., Shiao, H.-T., and Nelson, L. 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*, CHI '12, ACM (New York, NY, USA, 2012), 2811–2814.

[17] Yee, N., Ducheneaut, N., Yao, M., and Nelson, L. 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), 773–776.