Do Men Heal More When in Drag? Conflicting Identity Cues Between User and Avatar

Nick Yee[1], Nicolas Ducheneaut[1], Mike Yao[2], Les Nelson[1]

[1]Palo Alto Research Center
3333 Coyote Hill Road, Palo Alto, CA
[nyee, nicolas, lnelson]@parc.com

[2]Department of Media and Communication
City University of Hong Kong
mike.yao@cityu.edu.hk

ABSTRACT
Studies in the Proteus Effect have shown that users conform to stereotypes associated with their avatar's appearance. In this study, we used longitudinal behavioral data from 1,040 users in a virtual world to examine the behavioral outcome of conflicting gender cues between user and avatar. We found that virtual gender had a significant effect on in-game behaviors for both healing and player-vs-player activity.

Author Keywords
Avatar, identity, stereotypes, Proteus Effect, gender.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human Factors

INTRODUCTION

Digital environments allow us to create and customize our avatars (i.e., digital representations of ourselves) at the click of a button. Whether choosing to be an orc or an elf in a fantasy online game like *World of Warcraft*, or the exact shade of eye color in the user-created world of *Second Life*, virtual worlds (VWs) allow us to dramatically alter our self-representation. Understanding how people create avatars [2] and how social interaction is influenced by avatar appearance [10] are core research interests in the HCI community.

While it is easy to think of avatar creation as a one-way, user-directed process, studies have shown a more provocative two-way process between user and avatar. Specifically, the Proteus Effect [9] describes the observed phenomenon where users conform to expected behaviors and attitudes associated with their avatar's appearance. For example, users given taller avatars negotiate more aggressively in a bargaining task than users given shorter avatars. But what happens when an avatar's identity is in direct conflict with the user's identity? Given that avatar creation is typically a user-driven choice in VWs, it is difficult to know whether findings from randomly assigned avatars in laboratory studies generalize to natural settings.

In this paper, we describe a study that uses longitudinal data from actual users in a VW to examine the Proteus Effect when user and avatar identity cues are in conflict.

Outcomes of Conflicting Identity Cues

A conflict between physical and virtual identity cues may lead to a variety of plausible outcomes:

1) According to deindividuation theory, identity cues hold more sway when individuals are in VWs [4]. Thus, physical identity cues are deemphasized, and users conform to virtual cues, regardless of physical cues.

2) On the other hand, when virtual identity cues are made salient (especially when they are in direct conflict with the physical identity of a user), it allows for conscious monitoring and filtering of stereotypes [1]. Thus, users may not be influenced by the virtual cues at all. This seems particularly plausible in natural settings where users choose their own avatars.

3) But in more extreme cases, conscious monitoring may lead to overcompensation. This is similar to the ironic nature of mental control [5], such as asking someone to not think of a pink elephant. In much the same way, asking someone to not be aggressive may cause them to overcompensate and become overly docile.

A recent study [3] using longitudinal data from *EverQuest II* (EQ2) examined this conflict between players who do and do not gender-bend (i.e., play a character of the opposite gender). The activities they analyzed were combat, quest completion, player vs. player, and chatting. Among men, they found no differences between players who do and do not gender-bend. Among women, they found a mixture of overcompensation and deindividuation effects. In short, these findings lend some support to all three outcomes listed above.

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

We seek to expand upon these prior findings in two ways. First, we want to identify actual game-related gender stereotypes held by gamers. In the above-mentioned study, stereotypes from the physical world were assumed to transfer into a fantasy game context. As the authors in that paper noted, this translation resulted in broad measures of combat and questing that may not have been sensitive enough to detect gendered differences. Secondly, additional data from another online gaming context may help us more clearly see the outcome of conflicting identity cues. To address these two points, we gathered and analyzed stereotype and longitudinal behavioral data from the online game *World of Warcraft* (WoW).

METHOD

In order to identify stereotypically gendered behaviors to further analyze, we first conducted a web-based survey to explore which in-game behaviors were most strongly stereotyped in terms of gender. Below, we briefly describe the context of WoW. Then, we report the findings of the pretest survey of stereotypes and the analyses of in-game behaviors.

World of Warcraft

WoW is currently one of the more popular online games available commercially, with roughly 11 million active subscribers world-wide [6]. WoW uses a typical "leveling up" formula seen in computer role-playing games. The game encourages players in different ways to collaborate with other players. For example, users can create characters with different skill sets that complement each other (i.e., classes). Heavily-armored *tank* classes shield the group from enemy attacks while lightly-armored damage dealing *DPS* (damage per second) classes deal damage to enemies and *healing* classes restore health lost in combat. Player vs. Player (PvP) combat is allowed, but a set of rules defines where and whether players can attack and kill each other. And finally, it is important to note that in WoW, male and female avatars have the same abilities and skills. Thus avatar gender is entirely cosmetic and does not lead to any functional difference in terms of game-play.

Pretest of Player Stereotypes

Given that the Proteus Effect operates on shared social stereotypes, it is important to first understand what gender stereotypes WoW players have. 252 participants were recruited from message boards and websites catering to WoW players. Participants were asked to rate a set of six activities in WoW along a 5-point rating scale of whether they believed men or women had a preference for that activity: -2 (women strongly prefer), 0 (no difference), and 2 (men strongly prefer). We show the means and standard deviations for each activity in Table 1.

A paired t-test revealed that the stereotype valence for Healing was significantly different from that of Crafting ($t[251] = -8.20$, $p < .001$), and thus would be significantly different from all other activities as well. Thus, for a stereotypically female activity, we selected Healing to further analyze. And for a stereotypically male activity, we selected PvP. We did not select Tanking for the male activity to analyze because it is much more difficult to create a metric for Tanking from the current database of player activities (the WoW Armory, described in more detail below).

Activity	Mean	SD
Tank	1.08	0.76
PvP	0.98	0.76
Melee DPS	0.88	0.73
Ranged DPS	-0.23	0.65
Crafting	-0.25	0.6
Healing	-0.7	0.68

Table 1. Means and standard deviations for gender stereotypes in WoW, ordered from strongest male preferred activity to strongest female preferred activity.

Participants

1,084 WoW players were included in the next phase of the study. We recruited participants from forums dedicated to WoW, publicity on popular gaming sites (e.g., WoW.com), word-of-mouth on social media like Twitter, and mailing lists from previous studies of WoW players. We note that due to human subjects regulations, minors were excluded from participating in the study. Nevertheless, we were still able to gather data from a very wide age range (18-65). The average age of our sample was 27.03 (SD = 8.21). 26% of participants were women. These demographic variables are mostly consistent with previous studies of online gamers [8].

Procedure

Participants began by completing a web-based survey that gathered their demographic information. Participants were also asked to list up to 6 WoW characters they were actively playing. Once these characters were in our database, an automated data collection system was activated. The system then launched a daily web scraper to gather these characters' profiles (large XML files) from the WoW Armory over a 5 month period.

The WoW Armory is a public database maintained by Blizzard, the developer of WoW, which contains hundreds of longitudinal game-play metrics of every character, including total career kills, deaths, equipment currently worn, and even the number of hugs (i.e., emotes) given to other players. The Armory is updated daily.

Measures

To create a metric for healing, we calculated the ratio of total healing done against total damage done for each participant's primary character (i.e., the character that had the most days played during the 5 month period, $M = 53.66$, $SD = 39.71$). Calculating a ratio of healing against damage circumvents the potential problem of comparing hardcore

players from casual players. In other words, it is important to remove the noise of players who have a high amount of cumulative healing due to playing many more hours of WoW than the average. Thus, the ratio focuses more directly on the concept of preference—how does the player allocate between healing and dealing damage on a particular character? We also note that healing as reported on the Armory includes some noise from passive healing abilities that are beyond a player's control, but that these weak passive heals are easily overshadowed by direct healing spells in dungeon and raiding contexts. Similar to this healing metric, to create a metric for PvP, we calculated the ratio of PvP-related kills against total kills.

RESULTS

We first report some descriptive statistics for base rates of gender-bending, and then present the results for the research question.

Gender-Bending Base Rates

In our sample of 1,084 participants, we had 281 women and 801 men. On average, they had 2.79 (SD = 1.51) characters. Among men, 53.3% had a character of the opposite gender. On average for men, 33.4% of their characters were of the opposite gender. Among women, 18.5% had at least one character of the opposite gender. On average for women, 9.1% of their characters were of the opposite gender. The same trends were true for the primary character. 7.5% of women's primary characters were male. 29.3% of men's primary characters were female. This disparity between men and women is consistent with past data [7]. Also consistent with this past data, gender-bending is positively correlated with age among men (r = .10, p < .01), but not women (r = .08, p = .16).

Healing Metric

We conducted an ANOVA using Player Gender and Character Gender as the factors and the Healing Ratio (of each participant's primary character) as the dependent variable. We removed any outliers that were more than 2 SDs away from the mean.

The main effect of Player Gender was not significant (F[1, 1024] = .20, p = .67). In terms of Healing Ratio in the game, men (M = .33, SE = .02) and women (M = .30, SE = .06) were quite similar.

The main effect of Character Gender was significant (F[1, 1024] = 5.13, p = .02). Female characters (M = .39, SE = .03) had a significantly higher Healing Ratio than male characters (M = .23, SE = .06). Post-hoc comparisons revealed that this was true for both men and women (p's < .05). See Figure 1. The interaction effect was not significant (F[1, 1024] = .67, p = .41).

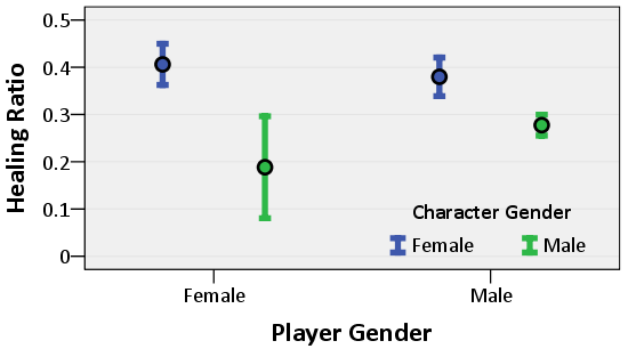

Figure 1. Plot of player gender and character gender interaction with standard error bars. Note that larger F-M variance in this and next graph are due to much smaller sample size of women who gender-bend.

PvP Metric

We conducted a similar ANOVA with the PvP metric. The main effect of Player Gender was approaching significance (F[1, 1096] = 3.29, p = .07). Men had a higher PvP Ratio (M = .06, SE = .002) than women (M = .04, SE = .007).

The main effect of Character Gender was significant (F[1, 1096] = 3.99, p = .04). Male characters had a higher PvP Ratio (M = .06, SE = .007) than female characters (M = .04, SE = .003). The interaction effect was not significant (F[1, 1096] = .00, p = .81). See Figure 2.

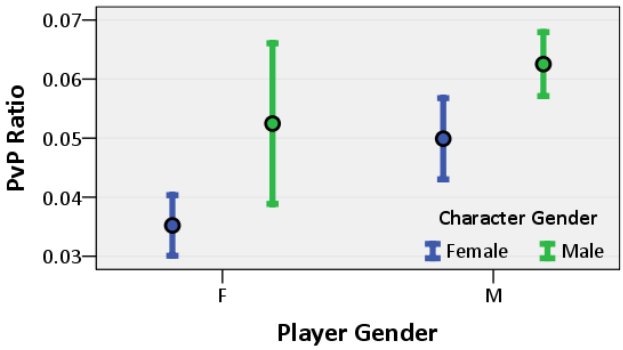

Figure 2. Plot of player gender and character gender interaction with standard error bars.

DISCUSSION

Our findings show that when identity cues conflict among actual VW users, users conform to their virtual identity. In terms of both healing and PvP activity, character gender had a significant effect on in-game behavior. This in turn suggests that deindividuation in virtual environments allows virtual cues to take precedence over physical cues.

On the other hand, even though there were some similarities, discrepancies with the earlier study of EQ2 data [3] complicate the picture. Similar to that study, we observed the Proteus Effect for PvP activity, but we found the effect for both men and women. And overall, we found more consistent support for the Proteus Effect in the current

data set, where both metrics showed a significant main effect for character gender.

We propose several reasons for these discrepancies. First, we conducted a survey to target gender-stereotyped in-game behaviors and likely identified more sensitive measures. Indeed, the PvP measure was largely comparable in both studies. Moreover, our survey results show that "combat" may be too broad as a metric, but finer-grained distinctions between Melee and Ranged DPS might show gendered differences. Secondly, different cultural norms and personality profiles among players in WoW and EQ2 may contribute to observed differences as well. For example, the EQ2 study found that 17.4% of men gender-bended, whereas in the WoW sample, 53.3% of men gender-bended. It is likely that another demographic or game-related variable is lurking as a moderating variable. In short, we can't take for granted that findings from one virtual world generalize to other virtual worlds, and we need to continue exploring identity cue conflicts in other virtual environments.

Social Construction of Gender

The data on healing presents a particularly provocative illustration of the social construction of gender. While WoW players share a strong stereotype of women preferring to heal, our findings show that this stereotype isn't true. We did not find a significant difference in terms of player gender. What we did find was that players enact this stereotype when gender-bending. Thus, the stereotype becomes true in the VW—female characters heal more than male characters. These findings illustrate how gender differences can be socially constructed, and how gender is being produced in VWs.

Implications for CHI

The Proteus Effect has implications for avatar-mediated systems in general, whether 2D or 3D. For example, these studies in the Proteus Effect suggest that providing more athletic and fit avatars in the context of a health game may lead to better outcomes than providing an avatar that resembles the user. Our findings suggest that any cue conflicts between user and avatar are likely to be resolved in favor of the avatar.

Limitations

There are significant trade-offs between lab experiments and natural observations. Given that character creation, and especially gender-bending, is a user choice, it makes sense to try to study this behavior in a natural setting instead of forcing users to adopt particular avatars. On the other hand, we lose experimental control in doing so.

This introduces several limitations to our study. First, it is possible that other players have a significant role in shaping a female avatar's behavior, via social pressure (i.e., behavior confirmation). It is also difficult to tease out whether an underlying personality variable is driving both gender-bending and healing preference. Nevertheless, our study illustrates the outcome of conflicting identity cues between user and avatar in an actual VW. While we cannot pin the causality precisely on the Proteus Effect, the findings are consistent with the theoretical prediction.

These limitations also point directly at possible future studies where avatar gender is incongruent between user and other players—i.e., a man playing a female character is perceived as a male character by other players. Such a setup would allow us to pit the Proteus Effect against behavioral confirmation directly.

Conclusion

Creating an avatar is a fascinating two-way process. Our findings in the current study suggest that we conform to the roles that are expected of our avatars, independent of who we are in the physical world. While one might assume that the infinite fluidity in VWs helps free us of our physical identities and norms, the opposite appears to be more accurate in this case. And in the particular case of healing, we in fact found that a gender stereotype that isn't true in the physical world becomes true in the virtual world.

ACKNOWLEDGMENTS

This research is sponsored by the Air Force Research Laboratory

REFERENCES

1. Devine, P. Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56* (1989), 5-18.
2. Ducheneaut, N., Wen, M., Yee, N. and Wadley, G. Body and mind: a study of avatar personalization in three virtual worlds. *Proceedings of CHI 2009* (2009).
3. Huh, S. and Williams, D. Dude looks like a lady: Gender swapping in an online game. in Bainbridge, W. ed. *Online Worlds: Convergence of the real and the virtual*, Springer, New York, 2009.
4. Johnson, R. and Downing, L. Deindividuation and valence of cues: Effects on prosocial and antisocial behavior. *Journal of Personality and Social Psychology*, *37* (1979), 1532-1538.
5. Wegner, D. Ironic processes of mental control. *Psychological Review*, *101* (1994), 34-52.
6. White, P. MMOGData: Charts. *2009* (2008).
7. Yee, N. The Demographics of Gender-Bending. *The Daedalus Project*. http://www.nickyee.com/daedalus/archives/000551.php
8. Yee, N. The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and Virtual Environments*, *15* (2006), 309-329.
9. Yee, N. and Bailenson, J. The Proteus Effect: The effect of transformed self-representation on behavior. *Human Communication Research*, *33* (2007), 271-290.
10. Yee, N., Bailenson, J.N. and Rickertsen, K. A Meta-Analysis of the Impact of the Inclusion and Realism of Human-Like Faces on User Experiences in Interfaces. *CHI 2007 Proceedings* (2007), 1-10.