Co-Constructing Virtual Identities: Insights from Linguistic Analysis

Ross Burkholder

University of Chicago
115 E 58th St. #224
Chicago, IL 60637
+001 231 215 4056
rossburk@uchicago.edu

ABSTRACT

This article critically examines the co-construction of personae for fictional characters in virtual environments. Expanding upon Gee's (2003) tripartite notion of identity in virtual worlds, this paper focuses on how virtual identities are created, and who does the creating. Using sociolinguistic methodology, I show how alterations in behavior based on avatar characteristics (The Proteus Effect: Yee and Bailenson 20007) can be used as a window into the virtual identity creation process. Potential contributions to virtual identity from three sources are analyzed: the community, the creators of the virtual environment, and influences from the non-virtual world, concluding that community created knowledge seems to play the most significant role in virtual identity construction.

Keywords

sociolinguistics, (virtual) identity, language, Proteus Effect, MOBA, avatars

INTRODUCTION

The relationship between a player and their avatar is one of the most crucial aspects of having a meaningful gaming experience. There is significant debate, however, on the details of this relationship. In his book, *My Avatar, My Self: Identity in Video Game Role-Playing*, Zach Waggoner makes use of the difference between an *Avatar* and an *Agent*. In this dichotomy, both avatars and agents are controllable representations of the player in the virtual environment. Avatars distinguish themselves in having a level of customizability, which agents lack (2009, 8-9).

In many ways this distinction reflects aspects of Gee's (2003) tripartite notion of identity in virtual worlds. In his book *What Video Games Have to Teach Us About Learning and Literacy,* Gee outlines a view of identity which includes three distinct identities. The real-world identity is the identity of the player of the game. This identity resides outside of the game, in a flesh-and-blood body, complete with all the experiences and knowledge garnered in the non-virtual world. The second identity is the virtual-identity, which is the identity of the representation or character that the player is controlling within the virtual environment. Lastly, Gee's projective identity is the combination of the first two identities. This is the hybrid identity of the player controlling the representation, a middle ground which is not entirely the players non-virtual identity, and not-entirely the identity of the character in game, but rather a blend of the two (Gee 2003).

Within this framing we can see agents as virtual characters which come loaded with a fully-fledged virtual identity. While the player gets to control these characters, and

often make important choices while doing so, there is nonetheless a sense in which some aspects of the characters are immutable. Lara Croft from *Tomb Raider* (Crystal Dynamics 2013) is an independent, capable adventurer, regardless of the skill or decisions of the player. Geralt of Rivia is gruff and battle-worn regardless of the player's decisions while parenting Ciri during the course of *The Witcher 3* (CD Projekt Red 2015).

An avatar, however, represents something close to a *tabula rasa* with regards to virtual identity, giving the player much more control over projective identity than they have with agents. The *Dovahkiin* from *The Elder Scrolls 5: Skyrim* (Bethesda Game Studios 2011), for example, is customizable in race, clothing, abilities, personality, allegiances, romances, etc…

This article examines virtual identities that fall somewhere in-between these definitions of avatar and agent. Characters in the video game genre of Multiplayer Online Battle Arenas (MOBAs) typically reflect a state of virtual identity which is more formed than the *tabula rasa* of Skyrim, but not as fully fleshed as well-known characters such as Lara Croft. Using data from the popular MOBA *Dota 2* (Valve Corporation 2013), the goal of this paper is to examine how these 'half-way' virtual identities are co-created using material from the game creator, knowledge from the non-virtual world, and crucially the community of players themselves. Furthermore, I ask the question of what and how much each source contributes to the virtual-identity.

OVERVIEW

In order to determine the effects and strength of each of the hypothesized sources, three tiers of variables are analyzed: sources of input, virtual character distinctions, and behavioral patterns.
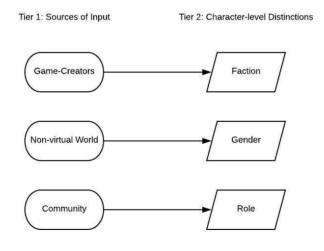

Figure 1: Each source of input connected to their corresponding character-level distinction.

Each of the three given sources of input (**game creators, non-virtual world, community**) critically informs our perceptions of one of the virtual character-level distinctions. For instance, material from the game creators determines which **faction** (Dire or Radiant, representing good and evil, respectively) a character is on, the non-virtual world influences our ideas of what **gender** means and how it is enacted, while the community itself, through the establishment of the meta-game, has control of which **role** each character plays in the game.

Each character-level distinction is in turn tested for behavioral patterns; specifically, differences in linguistic behavior. For example, all instances of players controlling male characters are grouped together, and the language data from these instances is processed looking for systematic differences. Each set of character distinctions is tested for linguistic indicators of **rudeness**, **politeness**, **tentativeness**, and **certainty**, using dictionaries based on those used in the Linguistic Inquiry and Word Count (LIWC; Pennebaker et al. 2015), a sentiment analysis tool. When significant differences are found in the linguistic behavior of a group, I hypothesize that these differences ultimately stem from the character-level distinction which defines the group. This hypothesis is in line with the Proteus Effect, which claims that the attributes of an avatar have effects on the behavior of the person controlling them (Yee & Bailenson 2007).
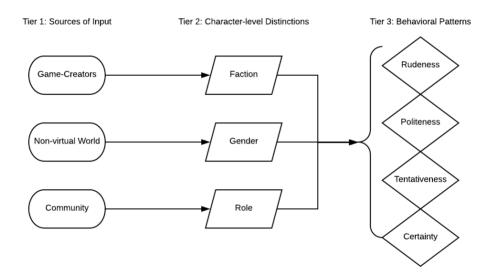

Figure 2: Each source of input connected to character-level distinctions and behavioral patterns.

Which character-level distinctions are linked to significant behavioral alteration gives an indication of which distinctions players are paying attention while controlling their avatars, and through this an indication of which sources of input have the greatest impact in the process of virtual identity creation. To give a concrete example, if it is found that female characters (character-level distinction) show a systematic bias towards politeness (behavioral alteration) in their speech, we can infer that the non-virtual world (source of input) is playing a role in the process of virtual identity creation.

DATA

The data for this study comes from the log data from 153,625 games of Dota 2, a popular online multiplayer game published by Valve Software in 2013. At the time of writing there are 115 unique characters that players select to control at the beginning of the game. Matches are divided into two teams of five players each. The object of the game is to work together to increase the power level of your team's characters, ultimately leading to victory through the destruction of the enemy team's base. A typical match of Dota 2 lasts between 30-60 minutes. Players can form a team of friends to play with, otherwise a matchmaking algorithm pairs players with other players of approximately their own skill level.

All data in the corpus is taken from the two North American Servers (NA East and NA West). The two most common languages on these servers are English and Spanish. No effort was made to divide the corpus by language for this project. The log data itself consists of game metadata (winner/loser, kills/deaths, heroes played) and the messages sent on the global chat channel. The resulting chat corpus contains ~7.5 million words.

BACKGROUND

This paper draws heavily from the traditions of psychology, sociolinguistics, and more recent works in computer-mediate communication (CMC), in addition to literature dealing with the nature of avatars and identity. In this background section, a brief overview of this literature is provided.

The Proteus Effect

The Proteus Effect was first introduced by Yee & Bailenson (2007) in a paper titled "The Proteus Effect: The effect of transformed self-representation on Behavior". In this seminal work, the authors run two experiments. The first experiment tested to see if controlling an attractive avatar affected a participant's behavior. It was found that participants controlling an attractive avatar divulged more information about themselves, and walked more closely to other interactants than their counterparts, who were given unattractive avatars (281-282). Attractiveness in avatars was determined based on the ratings of impartial judges. Note that the connection between attractiveness (avatar-distinction) and signs of confidence (behavioral-distinction) is one that has been pre-established in non-virtual environments (Snyder et al. 1977). Using the methodology from the present paper we could posit that the source of effect comes from the non-virtual world. Participants transfer their knowledge of what it means to behave like an attractive person from the non-virtual world and replicate these behavioral patterns in a novel virtual world.

In the second experiment, the authors tested the effects of avatar height (avatar-distinction), on behavior during a negotiation task. Overall, they found that participants controlling a tall avatar negotiated more forcefully than participants controlling shorter avatars (284-285). Again, this distinction mirrors observations made in non-virtual environments.

Peña et al. (2009) elaborates on Yee and Bailenson's (2007) Proteus Effect by running two additional experiments. In the first, participants who were given control of avatars in black robes demonstrated more aggressive behavior and less group cohesion than participants controlling avatars in white robes. In the second experiment, participants were given a virtual avatar and asked to perform a Thematic Apperception Test, writing stories about an image in the virtual world. Participants who were given avatars dressed in KKK-like clothes composed stories which

contained higher levels of aggression than participants controlling avatars dressed to resemble doctors. Here we see direct evidence of linguistic behavior being affected by avatar characteristics (Peña et al. 2009).

Psychological mechanisms

The psychological mechanisms which have been leveraged to explain these Proteus Effects have primarily been self-perception theory (Bem, 1972), the automaticity model of priming effects (Bargh & Chartrand, 1999), and Behavioral Confirmation (Snyder et al. 1977). Self-perception theory is the notion that people infer their own thoughts and beliefs from observing their own actions (Bem 1972). Thus, when a person is paid money for doing something that they already love to do, they may infer that they actually enjoy doing this activity less. The idea being that they observe themselves being paid to do it, and infer that the payment is the only reason for their behavior, or that the activity itself is one that people would not engage in without being paid.

With regards to avatars, players may infer from viewing the attributes of their own avatar, what kind of person they are. For instance, a player may observe themselves controlling an avatar that is tall and muscular, and then associate with themselves behavioral traits which they stereotypically associate with tall and muscular people (bravery, confidence, skill). This self-perception may then influence their behavior within the virtual world, producing the Proteus Effects observed in studies like Yee and Bailenson (2007) discussed earlier.

The automaticity model of priming effects has also been employed to explain the production of Proteus Effects (Peña 2011).

"...priming studies have shown how the perception of people, stereotypes, and situational cues influences the subsequent interpretation of target stimuli as well as perceivers' reactions to such stimuli." (Peña 2011, 152)

In this formulation, priming works through a 'spreading activation mechanism', where observation of a target stimulus activates a particular concept. This activation then spreads, making associated concepts easier to activate (Collins & Luftus 1975). The stimulus for priming can come in variety of forms, visual or auditory. For example, in a 1998 study by Dijksterhuis and Van Kippenberg, two groups were given a test of general knowledge. The first group was primed for 'intelligence' by researchers talking to them about professors before the test. The second group was primed for 'stupidity' by researchers talking to them about soccer hooligans before the test. The group primed for intelligence performed significantly better than the group primed for stupidity (Dijksterhuis and Van Kippenberg 1998). This experiment shows how in addition to concepts being more easily activated due to spreading activation, the priming of these concepts can be measured through *behavioral alteration*.

Many aspects of a virtual environment could be considered to prime interactants towards certain patterns of behavior, including avatars. One difficulty with priming effects in an uncontrolled setting is connecting behavioral alterations with particular primes. Any of the stimuli present in the virtual world could be priming a player, making them more or less likely to behave in a certain way.

Even when limiting the possible prime to a user's own avatar, it is often difficult to tease apart the effects of priming vs. self-perception. In the Peña et al. (2009) study, for example, it is difficult to tell if the participants were reflecting on themselves as

the type of people who would wear black robes, and thus acting more aggressively, or if they were primed by the color black or black robes, and that this activated concept was associated with aggression, leading to a change in behavior. It is not the goal of this paper to undertake this task of teasing apart the effects of these different mechanisms. Peña (2011) gives an interesting discussion on precisely this topic.

The core concept of behavioral confirmation is that when people go into a situation expecting to be confronted with certain behavioral patterns, they treat their interactant in a way that in fact produces the expected behavior. Snyder et al (1977) showed this effect in a simulated telephone conversation experiment. As part of this experiment male interactants held conversations with female interactants. The male participant was given a picture they were told represented the female participant. Half were given pictures of attractive women; the other half were given pictures of unattractive women. Attractiveness was determined by the ratings of independent judges. The female participants were kept blind to the study condition. The results of this study were that the female participants who were talking to men that thought they were attractive, responded in ways typical of attractive women, behaving friendlier and more confidently than the women interacting with men who assumed they were unattractive (Snyder et al. 1977). Thus, behavioral confirmation has an element of the self-fulfilling prophecy.

It has been hypothesized that players are not necessarily altering their behavior based on the appearance of their own avatar, but rather based on the way that other players in the virtual environment are treating them. The Proteus Effect, however, has been specifically outlined and tested to avoid this possibility. In the second experiment performed in Yee and Bailenson (2007), for example, the research assistant that the participant interacts with is kept blind to the height condition of the study, eliminating the possibility that the research assistant is treating the participants differently based on avatar attributes. In the second experiment of Peña et al. (2009), the participant performs the Thematic Apperception Test in isolation, eliminating the effects of behavioral confirmation. Given that our study takes place in a naturally occurring environment, however, such methods were impossible to employ in the current study and thus behavioral confirmation could play a role in producing the behavioral alterations that were observed.

Indexicality and Stereotypy

None of the psychological concepts cited as underlying the Proteus Effect provide a completely satisfactory explanation of the alteration in participant behavior. These mechanisms suggest *how* avatar-distinctions can result in behavior-distinctions, but it doesn't tell us anything about which pairing of avatar-distinctions and behavior-distinctions are actually made manifest. This paper concerns itself with exactly this information, for it is which connections are powerful enough to alter behavior that will tell us which aspects of the character's virtual identity players are cuing in to and being affected by. In this instance the *what* of these connections are tells us more about the process of virtual identity construction than the psychological mechanisms underpinning *how*.

Regardless of the which mechanism motivates the behavioral alteration observed as part of the Proteus Effect, they rely on the participant's underlying conceptual associations. Whether the player is acting more aggressively because they are primed by the black robes, perceive themselves, wearing black robes, as they type of person who would be aggressive, or whether they respond to people treating them like an aggressive person because they are wearing black robes, the connection between

black robes and *aggression,* or perhaps between *black* and *aggression*, is the crucial element of each mechanism. In order to formalize these connections, we turn to the semiotic notion of the *index*, and its extension *indexicality*, as used in both Linguistic Anthropology (Silverstein 2003) and Sociolinguistics (Eckert 2008).

In traditional Peircean semiotics, an index is a sign that has the power to index, or point to, conventionalized meanings (Peirce 1998). In linguistic anthropology and sociolinguistics much of the literature on indexicality concerns itself with non-referential indexicality. This tends to be when linguistic variables (different words or ways of speaking) are associated not only with the concrete or functional references which form the semantic structure of the utterance, but with socially created meaning, such as group membership. Thus, the apical realization of the final ING cluster in English (e.g. findin' vs finding), not only conveys its referential meaning, but also indexes the speaker as from the Southern United States (Campbell-Kibler 2007). Furthermore, this apical realization of ING might index the speaker as informal, less-educated, or hospitable, a cluster of social meanings and judgements which is associated with Southern-ness. This process of expanding associations is known as *indexical order* (Silverstein, 2003).
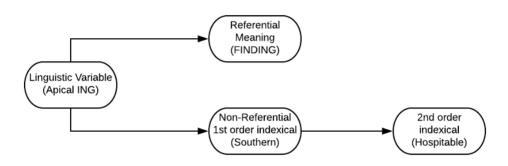

Figure 3: Flowchart showing the indexical associations of final apical ING clusters.

This process leads to linguistic variables becoming indexically associated with a whole range of possible associations, which has been called an *indexical field* by Eckert (2008). It is precisely these indexical connections which form the avatar-distinction to behavior-distinction pairings which we observe as part of the Proteus Effect. In the same way that the apical realization of the ING cluster indexes hospitality, black robes index aggression. Armed with the notion of indexicality, we can reframe the goal of this paper: discover what (and who) contributes to virtual identity by observing the indexical associations which players are paying attention to.

DESIGN & METHODS

The design of this study is as follows. First, avatar-distinctions were selected for each of the three sources identified as possible candidates for contribution to virtual identity: the game creators, the player community, and non-virtual influences. These avatar-distinctions were primarily chosen from the author's knowledge of the game and game community, supplemented by work from prior research where applicable.

For instance, gender was chosen as the avatar-distinction to represent influence from non-virtual sources, as there is a large body of existing literature which suggests that gender is an influencing factor on behavior both in non-virtual settings (Coates 1993; Herring 2003; Tannen 1994) as well as in other virtual settings (Baron 2004). If there

is a marked behavioral difference based on avatar gender, we hypothesize that this points to a strong influence on character personae from non-virtual sources.

The avatar-distinction chosen to represent influence from the game-creators has to do with the background lore given to each character by the creators. Within the virtual setting, each character is assigned to one of two factions. The map itself is split into two halves, with each representing one of the opposing factions, the *dire* and the *radiant*. The dire side represents a stereotype of villainy. Many of their characters are dressed in the colors black and red. Even their half of the map is dedicated to this trope, its landscape is black and lifeless, with dead trees and broken towers littering the field.

Figure 4: Shadow Fiend, a dire hero (left). The dire side of the map (right).

The radiant side is the opposite of the villainous dire, representing the heroes in the story. The heroes most often wear white, blue, and green, and their half of the map is lush, green, and vital.

Figure 5: Omniknight, a radiant hero (left). The radiant side of the map (right).

Thus, we hypothesize that if players are cuing into the hero's lore and history, we expect to see behavioral differences between radiant and dire characters. Here it is important to note the avatar-distinctions utilized by the game creators to establish characters as either being heroes or villains draws from material from the non-virtual world. The coloring of the characters is one major aspect of this, but others include the use of fire for villains, and light for heroes. While it is true that the difference between players drawing on their knowledge of gender-differences is not so different from them drawing on their knowledge of storybook heroes and villains, the difference in these categories is based on the explicit and intentional differentiation good/evil distinction on the part of the game creators, as opposed to any implicit gender differences.

Lastly, the avatar-distinction chosen to represent the input from the community of players are the in-game roles of the characters. Within Dota 2, each character has a unique set of stats and abilities which make them more or less suited for certain in-

game functions. Some characters, for example, deal large amounts of damage, while others have abilities which heal their teammates. Within a large spectrum of possible stat and ability configurations, semi-distinct roles have been carved out by the community into which characters fit. Often this role is based on what area of the map the character most typically plays. Of the three pathways (lanes) on the map characters which play in the middle lane are known as *mid* heroes. Those that play in the heavily forested spaces between lanes are known as *junglers*. The remaining three roles (*support, carry, off-lane)* face off against each other in the top and bottom lanes. Support characters use their abilities to heal and aid the carry they are paired with, carry characters deal large amounts of damage, while off-lane characters tend to have high health and are difficult to kill.

Though the characters tend towards these roles based on the stats and abilities they were given by the game-creators, it is nonetheless the community which ultimately determines the role of a character. Furthermore, it is the community which bestows upon each role a different attitude/personality (via indexicality), that will show through in player's altered behavior if they are indeed attuned to this avatar-distinction. The support and carry roles were chosen for this project specifically because they are seen by the community as existing in a complementary and binary relationship, which brings this feature into line with the other binary distinctions tested in this paper[1].

Testing Behavioral Alterations

In order to search for behavioral alterations, groupings of linguistic features were created. These groupings were based on the existing categories found in the Linguistic Inquiry and Word Count (LIWC) tool and the research which makes use of it (Pennebaker et al. 2015). Each category in LIWC provides a list of linguistic items which fall into the category, typically based on either the grammatical properties of words or the semantic/pragmatic meaning of the words. The *articles* dictionary, for example, is a simple list of the articles in English (*a, an, the*).

While the categories themselves are often simplistic, they have shown substantial ability to index particular social behaviors (Pennebaker et al. 2015). Newman et al. (2003), for instance, uses the categories *pronouns, motion verbs, and negative emotion words* to create a model which could predict truthful narratives from deceptive ones significantly above chance (62%), where human judges failed to do so (52%).

The unique environment from which the linguistic data for this project comes shapes the form of the data in ways that make using linguistic analysis tools, such as LIWC, challenging (Pennebaker et al. 2015). For instance, the length of an average message in our data set is unusually small, with each message composed of 2.25 words on average. This includes a large number of single word messages. This pattern likely stems from the fact that players are using the same input method to both chat and play, leading to a competition for resources that ultimately results in short messages, incomplete sentences, and frequent misspellings.

Non-standard spelling and meaning are additional issues when using generalized linguistic analysis tools. While some informal spellings are included in the LIWC dictionaries (ie *kinda* for *kind of*), some of the most frequent words in our corpus such as *lol* (laughing out loud), *gg* (good game), and *wp* (well played) are not included in relevant LIWC dictionaries[2] (Pennebaker et al. 2015). In addition, some words take on quite different meanings depending on the context. The word *easy*, for example, which may be considered a neutral or mildly positive word, is an insult in the context

of Dota 2, implying that the speaker's opponents were not skilled enough to make the game difficult for them.

With these reasons in mind, the dictionaries from LIWC were heavily modified to include frequent non-standard spellings and contextually specific meanings. Additionally, the dictionary names were changed. After these changes were made, four categories were selected for testing: rudeness, politeness, certainty, and tentativeness. Rudeness was designed to be the opposite of politeness, while certainty as designed as the opposite of tentativeness. The full LIWC dictionaries are proprietary, and thus not publicly available (Pennebaker et al. 2015).

Once these dictionaries were set, a simple automated token counting took place whereby each word in in the relevant sub-corpus was tested to see whether or not it was part of any of the four behavior dictionaries. There is no overlap between these dictionaries, meaning each word was tallied a maximum of one time. In order to ensure some degree of consistency, chat messages were only included for players who used 20 or more words per game. While this could feasibly skew the data towards one of the categories (players who speak more may tend to be angrier and thus ruder, for example), there is no a priori reason to think that this process should skew one of the sub-corpora more or differently from the rest. The percentages of words in each dictionary were then calculated and tested for significant differences when compared to the other category in their grouping (male vs. female, dire vs. radiant, support vs. carry). The results of this process are shown in Table 1.

The Benefits and Drawbacks of Naturally Occurring Environments

The majority of the work done on the Proteus Effect has been conducted in a laboratory environment, where control of environmental factors is significantly increased (though see Yee, Bailenson, Ducheneaut 2009 for a limited exception to this trend). In particular the authors have been careful to ensure that participants are only reacting to the appearance of their own avatar, and are not reacting to the way that others are treating them, as these reactions are also likely influenced by the avatar of the participant (Behavioral Confirmation; Synder et al. 1977). This control is typically accomplished by isolating the participant (Pena et al. 2009) or by ensuring that other interactants are blind to the condition of the study (Yee and Bailenson 2007), a control that is not possible in a naturally occurring environment like that of the present study.

Because the data for this study comes from the active community engaged in playing the game Dota 2, fewer controls could be put in place. There are at least two obvious limitations to the study as a result of this design choice. First, no demographic information could be collected about the players, as this was not available from the in-game data collected. While this lack of data does limit follow up studies, where it would be interesting to see if there were meaningful differences in the strength and direction of the Proteus Effects based on demographic distinctions, with the size of the data set from randomly selected games, we nonetheless assume that the study's sample is representative of the North American server player population.

A second limitation is that players were free to self-select their characters. The result of this lack of control is that it is difficult to tease apart the desired Proteus-like effects, wherein the aspects of the avatar cause differences in the behavior of the player, from the possible confound of players self-selecting avatars which reflect their pre-existing behavioral tendencies (aggressive people tend to pick characters with aggressive personae, and then act aggressively). While this is a serious concern in the present work, and prevents claims of causality between avatar-distinctions and

behavioral variation, the present study can nonetheless address the reality of the correlation that the players are confronted with and use this correlation as a window into the virtual identity creation process.

Furthermore, there are significant benefits to using data from a naturally occurring source. The most important among these is that the observed behavior is left in its naturally occurring context. In laboratory settings, participants are taken outside of a natural context, and put in an artificial one. While, as noted, this can have the benefit of reducing confounds, it also reduces the certainty that the behavior observed is reflective of something that happens in people's day to day lives outside of the laboratory. Additionally, by observing participants in a natural context the present study has access to locally constructed variables, such as character role, where much of the prior literature is forced to rely on larger order variables (such as gender), which are less effected by the oddity of the laboratory context.

On a similar note, laboratory experiments tend to be conducted on university campuses, primarily using university students as subjects. This tendency skews demographics towards younger, more affluent groups. Our naturally occurring data remedies this problem by taking a random sampling of the community in question. Overall, this paper follows prior literature (Yee et al. 2009) in determining that the benefits of naturally occurring context outweigh the drawbacks of having fewer controls.

RESULTS

The percentages for behavior distinction (column) by avatar distinction (row) are shown in Table 1. Each result box shows the percentage of words which were contained in the corresponding dictionary for the avatar-distinction shown. Thus, 3.427% of the words sent by players controlling male avatars were contained in the rudeness dictionary.

	Rudeness	Politeness	Certainty	Tentativeness
Male	3.427	0.881*	0.193	0.505*
Female	3.419	0.928*	0.192	0.536*
Carry	3.429*	0.877**	0.185**	0.496**
Support	3.336*	0.930**	0.207**	0.572**
Dire	3.454	0.888	0.196	0.514
Radiant	3.396	0.888	0.190	0.510

Table 1: Table showing the percentage of words contained in the LIWC dictionary (column) based on a sub-corpus containing chats from the relevant avatar-distinction (row). *p <.05 **p<.01

Overall, the differences in linguistic behavior between the categories of heroes appears small, but significant. This result is somewhat expected, as the influence of the Proteus Effect is subtle and subconscious. More than the differences in magnitude among the categories, the difference in which categories show significant

differences is enlightening. Of the four behavioral categories, all four show significant differences in the community determined role categories (carry vs. support), while the game-creator driven faction distinction shows no significant differences in any of the behavior categories. Influence from the non-virtual world, represented by the gender category, falls between the other two, with two of the four behavioral distinctions showing significant differences. These results point to a hierarchy of importance in the influences on the creation of virtual identity.

From the perspective of the behavioral categories themselves, there does seem to be some differentiation; rudeness and certainty show more and more significant differences than their counterparts rudeness and politeness. Whether or not this is a meaningful distinction, or instead an artifact of the dictionary selection/construction process is largely left to further research.

Post hoc Analysis: Individual Hero Results

Due to the relatively small effects seen in the primary test where characters are grouped by avatar attribute, an additional run of testing was done on each character individually. A benefit of this testing was to see if there were individual characters who did not follow the behavioral trend of their assigned category.

The results from the individual hero analysis showed larger differences than the group results. This is also an expected result, as grouping heroes together should have a smoothing effect on the behavioral differences shown among the individual heroes in each group. For example, 4.68% of the character Monkey King's (a male/dire/carry hero) words were within the rudeness dictionary, while only 3.13% of Abaddon's (another male/dire/carry hero) words were within this category. These individual hero results are more difficult to interpret than the grouped condition, since it is difficult to know what aspect of the hero underlies the observed behavioral pattern. They do, nonetheless, show variation by virtual representation, which lends additional support to the existence of the Proteus Effect in a naturally occurring environment.

DISCUSSION

The goal of this paper is to use the behavioral alterations in the language used by players controlling certain groups of heroes to shed light on the importance of various sources in the co-creation of a virtual character's virtual identity. A secondary goal of this paper is to provide evidence of the Proteus Effect in a naturally occurring environment.

Sources of Character Personae Construction

A driving question of this paper is who participates in the creation of the virtual identities for virtual characters. This paper hopes to raise questions of collaborative character creation, and to suggest that methodologies like the one used here are possible avenues for exploring the contributions of the multiple parties involved in this process.

While the findings for the group analysis showed weak behavioral alteration effects, the higher magnitude in variation and increased significance among the character role grouping suggests that players are paying more attention to community decided variables like character role (carry/support), rather than game-decided variables like a hero's team affiliation(dire/radiant).

This finding has important ramifications within the MOBA genre, as it provides evidence that game creators are not the only, nor even the primary, contributors to the perceived personae of the characters they create. From this perspective, the behavioral patterns observed in this project can be seen as a type of *participatory*

fandom, as written about by Jenkins (1992, 2003), wherein the consumers of this particular form of media can in turn become producers. Acknowledgement of this form of participatory fandom could in turn inform the differing possible reactions of how the game creators respond to and either choose to engage with, or ignore, the input and creations of their consumers (Shefrin 2004).

It may be the case that the increased variation found in the character role condition has more to do with the functional basis of the character role distinction, when compared to purely aesthetic variation, such as avatar gender, than the community-built aspect of this knowledge. The results of this paper could suggest that behavioral patterns are altered more readily in response to functional differences amongst avatars when compared to purely aesthetic ones. Differentiating between functional and aesthetic variables and their relevance to the Proteus effect was an unforeseen consequence of working with data within a natural context, as previous Proteus Effect research has been limited to studying the effects of avatar-based aesthetic variation on participant behavior. Rather than being an undesirable confound, these findings showcase a promising and important new avenue for future research.

The findings from the individual hero analysis may also suggest that the community contributes heavily to the creation of virtual character personae. One explanation for the increased variation among individual heroes is that the community assigns characters personality traits based on an individual basis, rather than being influenced by group trends among the avatars such as gender, team affiliation, or even role. The degree to which these individual virtual identities are influenced by game-created knowledge like character lore as opposed to community created knowledge is beyond the scope of the current work, however, and is left to future research.

Evidence for the Proteus Effect

Evidence for the Proteus Effect in the given data set is demonstrated by the linguistic analysis on both the grouped and individual hero conditions, with stronger support coming from the individual hero analysis, where heroes showed larger differences than the grouped condition.

The larger differences in the individual hero condition suggest that the personae of heroes were were obfuscated when different heroes were clustered together. The root cause of this phenomenon is less clear. It is possible that this result reflects that most meaningful avatar-distinctions which players are paying attention to were not selected, and thus a different grouping of heroes would show comparable results to the individual hero condition.

An alternate hypothesis is that the virtual identity of individual characters are co-constructed in a more individualistic manner, rather than reflecting a single or composite result of their avatar-distinctions. A character's virtual identity could be constructed on the basis of their individual attributes such as their unique skill set, voice lines, lore, or selection in professional tournaments.

The virtual character Sniper, for example, has the ability to simply click on an opposing character to deal extremely large amounts of damage. This damage is seen as not requiring a large degree of skill on the part of the player, resulting in the character being seen as a choice for beginners. In addition, Sniper can deal this damage at extremely long range, making it difficult for opponents to deal with. This simple and frustrating source of damage transfers its properties onto the character itself, with the result that Sniper is seen in turn as skill-less and annoying. This

process is taken one step further, transferring properties from the character to the player, resulting in the perception of people who play Sniper as annoying beginners.
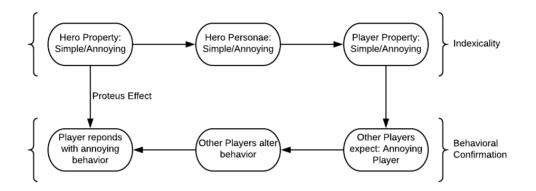

Figure 6: Flowchart showing the pathway from hero perception to player perception to player behavior.

Once these perceptions are established, we hypothesize that behavioral confirmation plays a role in shaping the behavioral patterns of the players controlling the hero (see Figure 6). Thus, the perception that Sniper is a no-skill frustrating hero, could actually cause players controlling Sniper to act in ways associated with unskilled and annoying players. This could explain Sniper's high level of recorded aggression, for instance (4.26% of words said by a player controlling Sniper were rude).

Regardless of these pathways, the existence of behavioral differentiation among heroes in our database lends supporting evidence for the Proteus Effect. This evidence is particularly important since the data comes from a large naturally occurring database. The addition of this support increases the applicability of the Proteus Effect to environments outside of the laboratory.

CONCLUSIONS

This paper hopes to contribute both to the sociolinguistics/psychology literature on the Proteus Effect (Yee and Bailenson, 2007), as well as to explore ideas of collaborative identity construction as found in game research (Gee 2003; Waggoner 2009). In straddling these two fields, I hope to show how the combination of quantitative and qualitative sociolinguistic methods can be of use in answering questions that go beyond the scope of traditional linguistic questions.

In this vein, the findings of this paper suggest that working backwards from observable behavioral differences, such as linguistic behavior, can provide a lens into the more abstract question of who gets to create a virtual character's identity, and to what degree? The three potential creative sources analyzed here were the game-creators, the community of players, and the non-virtual world. This paper is intended as only a beginning, however, and neither the sources investigated, nor the avatar- or behavior-distinctions used are intended to provide comprehensive coverage, but are rather an initial sampling of available sources/distinctions.

Possible avenues of future research include expanding on these categories, as well as employing similar methodologies on diverse and novel data sets. Each virtual environment comes with its own unique set of constraints, whether socially constructed or programmatically enforced (Herring 2007), and it is only by surveying a wide selection of environments and populations that we can avoid overgeneralizing the existing findings, and ideally uncover the patterns and processes which underpin them.

ACKNOWLEDGMENTS

BIBLIOGRAPHY

Bargh, J.A. and Chartrand, T.L. 1999. "The Unbearable Automaticity of Being." *American Psychologist.* 54, 462-479.

Baron, N. 2004. "See You Online: Gender Issues in College Student Use of Instant Messaging." *Journal of Language and Social Psychology*. 23(4), 397-423.

Bem, D. 1972. "Self Perception Theory." In *Advances in Experimental Social psychology.* 6, 2-57. *edited by Berkowitz, L.* New York, NY: Academic Press.

Bethesda Game Studios. 2011. *The Elder Scrolls V: Skyrim.* Windows. Bethesda Softworks.

Campbell-Kibler, K. 2007. "Accent, (ING) and the Social Logic of Listener Perceptions." *American Speech*. 82, 32-64

CD Projekt Red. 2015. *The Witcher 3*. Windows. CD Projekt.

Coates, J. 1993. *Women, Men, and Language* (2nd ed.) New York, NY: Longman.

Collins, A.M., and Loftus, E.F. 1975. "A Spreading Activation Theory of Semantic Processing. *Psychological Review*. 82, 407-428.

Crystal Dynamics. 2013. *Tomb Raider*. Windows. Square Enix.

Dijksterhuis, A. and van Knippenberg, A. 1998. "The Relation Between Perception and Behavior, or How to Win a Game of Trivial Pursuit." *Journal of Personality and Social Psychology.* 74. 865-877.

Eckert, P. 2008. "Variation and the Indexical Field." *Journal of Sociolinguistics.* 12, 453-476.

Gee, J.P. 2003. *What Video Games Have to Teach Us About Learning and Literacy.* New York, NY: Palgrave MacMillan.

Herring, S. 2003. "Gender and Power in On-line Communication." In *Designing for Virtual Communities in the Service of Learning,* edited by Barab, S., Kling, R. and Gray, J., 338-376. New York, NY: Cambridge University Press.

Herring, S. 2007. "A Faceted Classification Scheme for Computer-Mediated Discourse." *Language@Internet*. 4(1)

Jenkins, H. 1992. *Textual poachers: Television fans and participatory culture.* New York, NY: Routeledge, Chapman, and Hall.

Jenkins, H. 2003. Quentin Tarantino's *Star Wars*?: Digital Cinema, media, convergence, and participatory culture. In *Rethinking media change: The*

aesthetics of transition, edited by D. Thorburn and H. Jenkins, 281-312. Cambridge, MA: MIT Press.

Newman, M.L., Pennebaker, J.W., Berry, D.S., and Richards, J.M. Lying Words: Predicting Deception from Linguistic Styles. *Personality and Social Psychology Bulletin*. 29(5), 665-675.

Peirce, C.S. 1998. *The Essential Peirce*. Volume 2. Bloomington, IN: Indiana University Press.

Peña, J. 2011. "Integrating the Influence of Perceiving and Operating Avatars Under the Automaticity Model of Priming Effects." *Communication Theory*. 21, 150-168.

Peña, J., Hancock, J.T., and Merola, N.A. 2009. "Avatar Priming Effects in Virtual Settings." *Communication Research*. 36, 838-856.

Pennebaker, J.W., Boyd, R.L., Jordan, K., and Blackburn, K. 2015 *The Development and Psychometric Properties of LIWC2015*. Austin, TX: University of Texas at Austin.

Shefrin, E. 2004. "Lord of the Rings, Star Wars, and Participatory Fandom: Mapping New Congruencies between the Internet and Media Entertainment Culture." *Critical Studies in Media Communication*. 21(3), 261-281.

Silverstein, M. 2003. "Indexical Order and the Dialectics of Sociolinguistic Life." *Language and Communication*. 23(3-4), 193-229.

Snyder, M., Tanke, E. and Berschied, E. 1977. "Social Perception and Interpersonal Behavior: On the Self-Fulfilling Nature of Social Stereotypes." *Journal of Personality and Social Psychology*. 35(9), 656-666.

Tannen, D. 1994. *Gender and Discourse*. New York, NY: Oxford University Press

Valve Corporation. 2013. *Dota 2*. Online Game. Valve Corporation.

Waggoner, Zach. 2009. *My Avatar, My Self: Identity in Video Role-Playing Games.* Jefferson, NC, USA: McFarland & Company Inc.

Yee, N. and Bailenson, J. 2007. "The Proteus Effect: The effect of transformed self-representation on behavior." *Human Communication Research* 33, 271-290.

Yee, N., Bailenson, J., and Ducheneaut, N. 2009. "The Proteus Effect: Implications of Transformed Digital Self-Representation on Online and Offline Behavior." *Communication Research*. 36(2), 285-312.

ENDNOTES

1 The use of gender as a binary distinction in this paper is not intended to reflect the author's beliefs on gender politics, but rather in recognition that within the virtual world of Dota 2, stereotypical gender-binary representations are often used.

2 Many popular internet words, such as *lol* and emoticons, are included in a separate internet-centric dictionary, which is not of interest in the current project.