Avatars of a Feather Flock Together: Gender Homophily in Online Video Games Revealed via Exponential Random Graph Modeling

Shiyu Zhang, Sander Bakkes, Diederik Roijers, and Pieter Spronck

Abstract—**Increasingly more scholars have regarded the virtual worlds of massively multiplayer online games (MMOGs) as a social laboratory, and have paid research attention to the online interactions between its large number of players. In the present study, we examine a widely observed real-life phenomenon gender homophily (i.e., people of the same gender flocking together) in this virtual context. Specifically, we investigate how collaboration networks in an MMOG (*Destiny*) are shaped by the adopted gender of the avatar characters. This focus is interesting, as avatar gender in video games is generally a choice that is less constrained than it is in real life. To investigate the effect of gender in online video games, while controlling for the effects of several confounding factors, we employed a technique called *exponential random graph modeling*. Mirroring how interpersonal relationships are often gendered in real life, and despite common phenomenon such as gender swapping, we found evidence supporting gender homophily in the MMOG environment.**

Index Terms—**Exponential random graph (ERG) modeling, game analytics, gender homophily, online video games, social networks.**

I. Introduction

EVEN with the most advanced cosmetic and surgical methods, we cannot fully control how we present ourselves. Gender, age, and socio-cultural constructs unquestionably affect and shape our interpersonal interactions. In contrast to the constraints of real life, one may argue that nowhere is self-presentation more manageable and voluntary than virtual environments [1]. With vivid graphics and immersive stories, massively multiplayer online games (MMOGs) build virtual worlds, which attract a large number of players to construct avatars as representations of themselves in order to achieve, to experience, to interact, and to bond. Thousands and even millions of diverse players enter these games with different motivations [2]–[4], create avatars and interact with each other anonymously through their customized images; this is something that makes MMOGs an ideal laboratory for research [3].

Among various characteristics one can customize for his or her avatars, gender is one of the most important ones. Numerous features that may stratify a society (e.g., regarding socioeconomic or cultural factors), are not necessarily reflected in avatars. Still, however surreal MMOGs characters may be, a choice of gender is almost always available in *Destiny*, *World of Warcraft*, *EverQuest*, and *Eve Online*, to name a few. The salience of gender for avatars accords with the social realities that gender is one of the strongest categories for impression formation [5] and human interactions largely revolving around genders [6].

While the effect of gender in real-life interactions has been extensively studied [7], empirical evidence about this in virtual game environments has just started to accumulate, as commercial MMOGs have only been developed for less than three decades. Relatively little is known about how gender plays a role in interactions when people can freely express themselves or if gender is a relevant factor at all. To contribute to closing this gap, in the present study, we focus on the effect of gender on a specific form of interaction—*collaborative gameplay*—between avatars. People collaborate with one another in games not only because it is fun but also because it is an (often mandatory) part of the game experience. Motivated by recreational reasons and by the need of relying on others expertise, in-game collaborative relationships are central to the social aspect of MMOGs [8].

Bridging game analytics, network analyses, and socio-psychological studies on how human interactions/relationships are influenced by gender, we examine how collaboration networks in an MMOG were shaped by avatars gender (research question). In what follows, first we outline how relationships and interactions are gendered in real life; then we discuss why a study on avatar gender in the context of MMOGs may add a unique perspective to the existing real-life observations; next, we review a few studies that investigated networks or gender-related questions in the MMOG environments; and finally we explain why applying methods of network analyses is suitable to the present study and how it can contribute to research in the field of game analytics.

Manuscript received October 31, 2017; revised January 28, 2019; accepted November 18, 2019. Date of publication December 23, 2019; date of current version March 17, 2020. *(Corresponding author: Sander Bakkes.)*

S. Zhang is with the Department of Cognitive Science and Artificial Intelligence, Tilburg University, 5037 Tilburg, The Netherlands, and also with the University of Michigan, Ann Arbor, MI 48109 USA (e-mail: zsy@umich.edu).

S. Bakkes is with the Department of Information and Computing Sciences, Utrecht University, 3512 JE Utrecht, The Netherlands (e-mail: s.c.j.bakkes@uu.nl).

D. Roijers is with the HU University of Applied Sciences Utrecht and the AI Research Group, Vrije Universiteit Brussel, 1050 Ixelles, Belgium (e-mail: diederik.yamamoto-roijers@hu.nl).

P. Spronck is with the Department of Cognitive Science and Artificial Intelligence, Tilburg University, 5037 Tilburg, The Netherlands (e-mail: p.spronck@tilburguniversity.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TG.2019.2961836

II. Related Work

The old proverb says that "birds of a feather flock together"; I suppose that...similarity begets friendships. (Plato, 360 B.C.E)

A. Gender Homophily in Real Life

The phenomenon that people of similar characteristics flock together is referred to as *homophily*. Across the entire lifespan, friendships are established more often between the same-gender people; such a gender homophily pattern has been robustly observed in various cultural contexts for many decades, e.g., [6], [7], [9]–[13]. Children express a preference for the same-sex friends very early on [7], [14] and this preference reaches a high degree during their primary school years [15]–[17]. Although friendships with the opposite gender become more common as children enter late adolescence [17]–[19], the opposite-gender friends generally do not supplant the same-gender ones [19], [20]. Continuing this pattern, the same-gender relationships remain to be a major part of friend networks for young [21]–[23], middle-aged, and older adults [24].

Gender homophily and gender segregation are two sides of the same coin; a line of research focusing on gender segregation and the accompanying wage gap in workplaces has revealed a bitterer side of such a tendency. Male and female are over-represented in different school majors and subsequently in different occupations [25]–[27]. The gender segregation is often associated with men being in more profitable or prestigious positions than women [27], [28]. The occupational gender segregation, which certainly results in the homophily of peoples interpersonal networks [6], seems to also result from the gendered networks [29], [30] because social capital gathers and flows through men‘s and women's networks unevenly [30]–[32]. The crux of the empirical evidence is that people's interactions and relationships in various context are shaped by gender.

Different reasons have been proposed to explain gender homophily, including personal preferences, social pressure, social norms, and physical (spatial) constraints [7], [14]. First, people initiate interactions and maintain relationships with the same-gender others because of their psychological preference toward the others that are alike themselves [14], [33]. Second, gender homophily can also indirectly result from differences in activity preferences [7], [14]. For example, previous research suggested that women might engage in more emotional sharing and talking with friends, whereas men might prefer actively doing things together [34]. Third, peers, parents, teachers, and, later, romantic partners may encourage same-gender friendships because these are considered more appropriate [7], [14], [35]. Fourth, social norms stereotype which are considered legitimate and qualified in what roles and this is a crucial reason for gender segregation in formal settings such as education and labor market [32], [36]. Fifth, spatial propinquity (i.e., the physical closeness) between people also influences friendship formation [37]. For example, the separation of school majors and occupations result in women being surrounded by other women and men by other men [26], [27], which may cultivate same-gender friendships.

B. Avatars in MMOGs

The current focus on collaborations in MMOGs shares similarity with both real-life friendships and colleague relationships. On the one hand, given the recreational nature of MMOGs, it is reasonable to categorize in-game relationships as friendships. On the other hand, given the competitiveness of the games, players may choose to game with others who are sufficiently competent and share similar goals and, thus, in-game collaborations hold a connotation of colleague relationships. How will gender play a role in such collaborative relationships? We suggest that an answer to this question in the MMOG context which may contribute to research that was mostly based on the real-life society from at least two perspectives.

For one thing, in regard to how gender influences in-game collaborations, a comparison between the real life and game worlds generates interesting insights because of some above-mentioned reasons that encourage gender homophily in real life that become less salient in virtual environments. To start, acknowledging that players enter MMOGs with different motivations [3], [4], it is reasonable to state that most of them share the same purpose playing the game and, thus, incompatibility in activity interest is less relevant for shaping gamer's interactions. Next, given the relative anonymity of virtual worlds, people have the opportunity to manage an identity and establish a network free from gazes in their real-life milieu [2], [3], [38], and this should largely relieve them from the social pressure of their surrounding people. Furthermore, although the gender stereotypes remain prevalent in the game context, for example, females are often expected to be feminine and adopt the supporting roles (e.g., a healer; [39], these stereotypes may not directly hinder collaborations between males and females because character skills are often complementary and, generally, different skills are required for winning combats. Finally, virtual interactions transcend the constraints of physical space. Although people who are spatially very far away (e.g., different countries) may have less chance of interacting because, for instance, they use different servers, this should not be a problem when the focus is interactions between genders. Removing or weakening various effects, it seems that preferences are left to be the most important reason. Studies about networks in the virtual environment like the current one may contribute to teasing the effect of psychological preferences apart from other factors.

For another, since players can construct avatars as a part of their virtual identities and all interactions are mediated by avatars, an examination of avatar gender offers a unique opportunity to probe how the concept of gender influences relationship formation. Players can either express their real-life gender on the avatars or experiment with the opposite identity. This flexibility of online self-presentation makes the examination of avatar gender attracting but also brings challenges to clearly interpreting the concept of gender. To emphasize this attraction and address this challenge, we suggest that the meanings of gender in the actual society seem to be carried over to the virtual worlds onto the avatars. Following the previous research, we argue that avatars are not value-free images, the choice between a male or female character is often motivated by meaningful reasons, and the avatar gender may, in turn, influence how masculine or feminine the players operate their avatars in game.

Before zooming in on the meaning of avatar gender, we first discuss the broader background of MMOGs and emphasize that, although MMOGs are marked as games, the social and psychological aspects of them are not trivial [3]. Once they are launched, the participation of thousands and even millions of players makes the game an evolving and unending collective drama [40, p. 6]. Empirical studies have found that players engage in games more than 20 h per week on average and that many of them disclose their secrets to their virtual friends, lean on each other for emotional support, and even experience romantic clicks with other players [2], [3], [41]. These experiences are mediated by avatars; no wonder that players identify and attach with their avatars [42], [43]. However, we wish to mention that there are different subcategories within MMOGs and the importance of avatars might not be the same in them. For example, some games are defined as massively multiplayer online role-playing game (MMORPG). As its name implied, avatars and role-playing are important elements of these game settings. World of Warcraft and Everquest are examples of this type of games. Some others are defined as massively multiplayer online first-person shooter game (MMOFPSG). These games offer a first-person perspective and the players often do not see the images of their avatars in combat (though they can still see other players avatars). In games like this, players might not pay as much attention to their avatars as they do in MMORPGs. The focus of this study (Destiny) is more of the latter type (MMOFPSG). Below, we cited some studies based on the former type MMORPG (e.g., World of Warcraft and Everquest) to help framing our theoretical arguments because these studies shed light on the social aspects of gaming that we are interested in, but we emphasize the differences between the different kinds of games and that findings on the MMORPGs might not be closely applicable to the settings of MMOFPSG.

Specific to the choice of gender for avatars, one important evidence, supporting that such choices are not random, hinges on the empirical findings that a majority of players maintain their real-life gender in game environments. A large-scale survey conducted on fan sites found that around 85% of EverQuest players make their avatar genders the same as their own [41]. One research using data from Fairland Online, in which players can construct an unlimited number of avatars, found that two-thirds of players owned only same-gender avatar(s) [44]. One other study on World of Warcraft players found a strong correlation between player gender and avatar gender ($r = .7$) [45].

One may reasonably point out that, although the majority players keep their actual gender, the amount of players who operate opposite-gender avatars is surely not negligible. Acknowledging this fact, however, previous research that looked into players motivations behind these so-called gender swapping actions seemed to support, instead of repudiate, that in-game gender bears meaning. The following quotes cited from two studies, which collected open-ended responses [46, p. 50] and [47, p. 94], provide some insights into players motivations of gender swapping. One common reason of gender swapping is the curiosity to “see if it felt any different.” An opposite-gender avatar “enables me to play around with aspects of my characters that are not normally easy to experiment with in real life” and a male avatar for a woman is “kind of like a window into their strange man universe.” Also, some women play with male avatars to disguise their female identity because when on a female [character], men think you do not know how to play, cannot be hardcore, and try to give you things to hit on you and she “was tired of creepy guys hitting on my female characters.” Interestingly, one man‘s trash is another man’s treasure. Some men choose female avatars because “if you make your character a woman, men tend to treat you FAR better” and “if you play a chick and know what the usual nerd wants to read, you will get free items.” These quotes certainly cannot represent all motivations behind gender swapping, but they suggested that there might be a general perception about what male and female avatars represent.

The meaning of real-life gender is further carried over to avatars implicitly because navigating in games with a male (or female) avatar may unconsciously bias people’s behaviors to be more masculine (or feminine), regardless of the actual gender of the player. This is demonstrated by a line of research studying how virtual avatars change people’s behaviors [1], [48], [49]. One possibility for such influence of avatars rests on behavioral confirmation [50]. The appearance of a player’s avatar affects how it is perceived and treated by the surrounding people; the treatments are fed back to the player and, in turn, affect his/her behaviors. In this regard, a player operates the avatar to be more manly because others expect his/her male avatar to act like a man [48]. Another possibility is referred to as *Proteus effect* [1], [51]: In anonymous contexts, people’s actions are directly shaped by the identity cues conveyed by their avatars. For example, a study found that, while men are often reluctant to seek help, using female avatars seemed to relieve them from such inhibition [48]. Similarly, another study found that women exhibited stereotypical masculine traits when they controlled male rather than female avatars [44].

To summarize, we aim at demonstrating that studying avatar behaviors in MMOGs add unique perspectives to real-life research on how interactions are shaped by gender. In the virtual environment, people’s preferences are minimally restricted; people can either maintain their true gender identities or explore the opposite possibility. While the fact that some players engage in gender swapping inevitably leads to doubts about the meaningfulness of avatar gender, we argue that the concept of avatar gender inherits meaning from the real-life society. In this regard, concerning our research question of how avatar gender influences collaborative relationships, we hypothesize that, just like people of the same gender flock together in real life [6], [7],

> *avatars of the same gender may collaborate more with each other (i.e., gender homophily).*

C. Empirical Studies About MMOGs

The social nature of MMOGs has attracted recent research attention. Some of these studies relied on surveys of game players, whereas others have started to use data from the games themselves enabling large-scale research on behavioral records.

The research by Ducheneaut *et al.* [52] was one of the first studies like this. They analyzed *World of Warcraft* data, graphed the social network for a typical guild (community), and examined the impact of guild memberships on playtime and leveling time. Pirker *et al.* [53] focused on the game *Destiny* and investigated whether players' network embeddedness was associated with their performance. Again in the context of *Destiny*, Rattinger *et al.* [54] combined behavioral profiling and social network modeling and tested whether players with the same play style tended to gather together in social networks.

Some studies specifically addressed gender-related questions. For example, Williams *et al.* [55] showed that male and female *EverQuest II* players differed in their motivation, enjoyment, and health, as well as that the behaviors of the two genders affirmed gender roles. Cole and Griffiths [2] surveyed players of various MMOGs and investigated how male and female players differed in their experiences of online friendship. Using data from Pardus, Szell and Thurner [56] examined how female avatars' networks were differently organized from male avatars. They found that females had more communication partners, attracted more positive behaviors, reciprocated friendships more; interestingly, they also found evidence for female homophily in trading activities.

While the previous studies offer valuable insight into the social aspects of MMOGs, one thing worth noticing is that many analyses on in-game social networks portrayed the networks based on one characteristic at a time, but seldom probed various factors simultaneously and investigate their relative importance to the network formation. However, looking at several factors is important to understand the forming mechanisms behind the observed networks. For example, as in the current study, though we hypothesize that gender is relevant to collaborative relationships, this cannot be the only reason why people team with others. To untangle the effect of gender from confounders, we perform social network analysis to examine how gender influence collaborative relationships, while controlling for several relevant factors, as presented in the next section.

D. Applying Network Analyses on Collaboration Networks

Besides gender, there are many more reasons why people interact and relate with one another; network analyses on the pattern of ties in social networks enable inferences on tie formation process. For instance, assuming that gender homophily is indeed a relevant mechanism, then ties in collaboration networks will cluster on the basis of gender, and a model that assigns a high probability to same-gender ties would fit the observed network well. By modeling probabilities, several tie formation mechanisms can be tested simultaneously for their relative importance. This is the principle of exponential random graph (ERG) modeling. To isolate the effect of avatar gender on collaborative relationships from other mechanisms, methodologically, we adopted the ERG modeling (refer to Section III and the Appendix for discussions that contrast ERG modeling and logistic regression); theoretically, we followed the framework discussed by Wimmer and Lewis [57], which outlined several reasons behind tie generations and we controlled for three types of confounders that may influence the MMOG network composition:

First, *propinquity* refers to the closeness induced by sharing physical space or activities. This is relevant to network analysis because people who are geographical distanced (i.e., low in propinquity) are less likely to form relationships [37]. Although networks in virtual worlds transcend physical constraints, a relevant definition of propinquity in MMOGs is time. If the active times of two players do not overlap, then it is impossible for them to collaborate. Thus, we control for overlap in playtime as an indicator of propinquity.

Second, the effect of one characteristic can be driven by the intersection of this characteristic with other characteristics [57], [58]. As a hypothetical example of this, if all male avatars are at performance level m and all female avatars are level n, and if different levels do not collaborate, then we will observe a perfect gender homophily, which is, in fact, driven by level homophily. To control for the potential effect of characteristic intersection, we include several basic avatar characteristics that we managed to gather, including avatar class, performance level, and the number of combats they engaged in on a single day (for a detailed explanation, please refer to the Section III).

Third, one other set of mechanisms that is important to tie formations is overarchingly referred to as endogenous networking mechanisms [59]. These mechanisms model the phenomenon that ties in a network are interdependently formed [58]. One of such mechanisms is triangle closure (or transitive triads; [57]–[59]: if A and B share a common friend C, then A and B are more likely to befriend each other. This tendency can amplify the existing imbalance in the network composition. Another mechanism, named preferential attachment, concerns the phenomenon that a popular person tends to become more popular because he/she is more visible (i.e., the rich get richer) [59]. This tendency can result in a highly skewed network with a small proportion of people being particularly well connected. Both mechanisms are relevant in MMOGs: one is probably more likely to game with a teammates teammates; a highly active avatar star can be a more attractive collaboration partner. As such, we control for both triangle closure and preferential attachment.

One study by Huang *et al.* [60] applied the method of ERG modeling to a network of 1525 *EverQuest II* players and examined how social, temporal, and spatial propinquity explain network ties between players. Gender homophily was part of their investigation (social propinquity) and they found that gender was not a factor influencing probability of tie formation. Our study was different from theirs in at least three aspects. First, their data were based on one server targeting North Americans whereas our data were not restricted to one server. Second, *EverQuest II* is a role-playing game and their context was player-versus-environment (PvE). In contrast, *Destiny* is a first-person shooting game and our context was player-versus-player (PvP). Our setting probably targets more competitive gaming and, thus, collaborative relationships in our study are likely to have different meaning. Third, their network was based on player and our network was based on avatars, thus we focused on a different aspect of gender comparing to their study.

Before ending this section, it is worth discussing the meaning of avatars collaboration networks in the current study. How to define meaningful relationships between avatars is a common challenge in in-game network studies [53]. Some research relied on explicit in-game concepts to identify relationships [61], such as groups/clans/guilds that players join [62] and friends or enemies directly marked by players [56]. Other studies looked for implicit relationships that can be derived from in-game interactions [61], such as from combat histories [53], [54], [61], colocation (playing in the same zone in the game) with guild members [52], and trading activities [56], [63]. We follow the latter approach and constructed collaborative relationships from avatars combat histories: if a pair of avatars gamed together in at least three combats on one day, then we drew a collaborative tie between them.

Fig. 1. Example of female and male avatars in *Destiny*.

E. Present Study

In summary, the present study investigates how collaboration networks among avatars in an online video game are shaped by the chosen gender of the avatar characters. Given that gender homophily is robustly observed in real life and that avatar gender seems to carry similar meanings as real-life gender, we hypothesize a pattern of gender homophily in the collaboration networks. That is, there might be more collaborations between avatars of the same gender than of different genders. To test the hypothesis, we perform statistical modeling of network analyses and control for various mechanisms that might confound the effect of gender on tie formation.

III. Method

The present study is based on data of 6376 *Destiny* avatars. In this section, we first provide a background introduction to the game *Destiny*, then explain how the data set was sampled using snowball sampling method, and finally introduce the statistical procedure for analyzing the sample.

A. Background of the MMOG Destiny

Destiny is an online-only multiplayer first-person shooter video game. Avatars in *Destiny* take on the role of “guardians,” whose goal is to defend the earths last safe city against alien races. The game is available on PlayStation 3, PlayStation4, Xbox 360, and Xbox One consoles. Players on a given type of console can only interact with other players who are on the same console, and, thus, the player networks for the four consoles are separate from each other. The current study focuses on players on the Xbox console, though differentiation between Xbox 360 and Xbox One cannot be made based on the available information.

Activities in *Destiny* are divided into PvE and PvP types; the former consists of story missions and the latter consists of combats between two teams of players. These two types of setting are separate and players performance rankings in the two settings are independent of each other. PvP activities are the focus of this study.

Players can choose their guardian avatars from three classes: Titan, Warlock, and Hunter. The three classes differ in skill and expertise. For example, Titans are strong in close combat, Warlocks have high offensive and recovery abilities, and Hunters focus on agility and mobility. Players can choose a gender for their avatars and customize their appearance (e.g., hair style, skin, and eye color). Fig. 1 gives an example of male and female avatars in *Destiny*. Unlike classes, male and female avatars do not differ in abilities. Each player can create at most three guardian avatars, one from each class. These avatars can differ in their gender setting. The current analyses are at the avatar level and, therefore, some of the analyzed avatars were controlled by the same players. The difference between in-game avatars and real-life players is worth noticing and in this study, the word “avatar” (“guardian avatar”) and “player” are used distinctively.

B. Data Set

The data were web-scraped from a PvP leaderboard[1] (i.e., a ranking board) of *Destiny*. The leaderboard only contains players who partake in ranked combat and these players probably play more competitively. The ranking is based on combat rating provided by the developer of *Destiny*; higher-performed players have higher ratings. The Xbox version of the leaderboard contained information on almost 2.8 million players; as explained below, only a small portion of the data was sampled. Detailed data of players can be found on the leaderboard, including the avatars each player created, the setting of each of the avatars, and the play history of the avatars. Under each avatars play history, data on each combat is available, including the starting time of the combat, the combat duration, and whom the avatar teams with in the combat.

In each combat, players can either intentionally team with certain people (premade group) or let the system to match them with random strangers (pick-up group). Only the intentional relationships (premade groups) were of interest to us, but we did not know whether the players intentionally chose their teammate or were assigned to game with stranger. However, given the large size of the avatar pool, the more a pair of avatars played together,

[1]We scraped our data set from the website http://pvp.planetdestiny.com/leaderboard/xbox/all-time/all-pvp. But a new version of the game (*Density 2*) has been released and statistics on *Destiny* that we used are no longer publicly available.

the less likely that their relationship was a result of the system random matching. Therefore, we leveraged the information on the number of combat between each pair of avatars to extract (potential) intentional relationships: we aggregated the shared combats between avatar pairs over a day; if a pair of avatars gamed together in at least three combats, then we assumed that the relationship between them are intentional and voluntary. We set a one-day period over which the combats were aggregated so that each avatar did not collaborate with too many partners and, thus, the network size remained manageable. The data were sampled from activities on April 8 (Saturday), 2017.

To draw a sample from the ranked avatar population and construct the collaboration networks among them, we adopted a link-tracing method that is analogous to snowball sampling (also known as chain referral method) [59] but more reliable. Different from regular snowball sampling, we did not rely on respondents to refer or recruit eligible subjects; instead, we extracted network ties from stored data and were able to construct a complete network for the avatars in our sample. We refer to our method as snowball sampling for the simplicity and intuition. Starting from a random avatar on the leader board (the initial seed), the network crawler written for the current study found the initial seeds collaboration partners (wave 1), include these avatars in the data set, continued to find the wave 1 avatars collaboration partners (wave 2), and so on until the size of the network (without repetition) reaches 5000. After 18 waves of snowballing, a network of size 6376 avatars[2] was constructed.[3] Fig. 2 graphically presents the method of snowball sampling.

The leaderboard does not contain personal information of the players, so we gathered data only on avatars, including their gender, class, and play history. Afterward, we calculated the number of combat encounters each avatar engaged in and avatars active playtimes based on their play histories. The features we collected or calculated are presented as below.

1) *Gender:* Avatars have the gender of either male or female. In all, 29% of the sampled avatars were female.
2) *Class:* In the current sample, 31% of the avatars were Titans, 29% were Warlocks, and 40% were Hunters.
3) *Performance:* We used an avatar's kills/deaths ratio as an indicator of its performance. A higher kills/death ratio indicates better performance. The score ranged from 0.13 to 2.92 in the current sample. This variable was standardized for the analyses.
4) *Number of Combat Encounters:* We calculated the number of all combats each avatar played on April 8, including the ones that they played with actors outside the current sample. On average, each avatar played 13.6 combats.

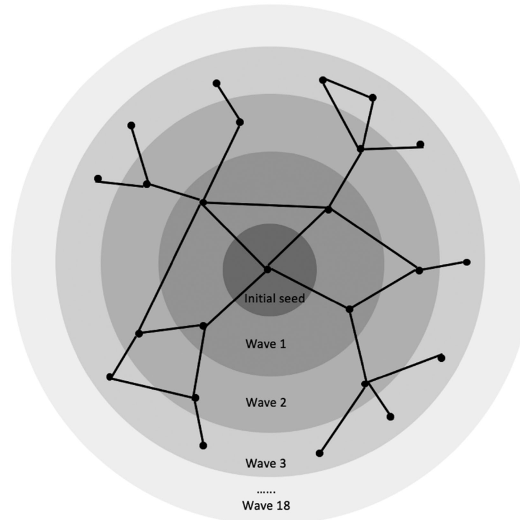


Fig. 2. Example of snowball sampling. Starting from an initial seed character, the network expands for 18 waves.

5) *Playtime:* Relying on the 24-h clock, we counted avatars' playtime based on the starting time and duration of the combats they played. For example, if an avatar started a 9-min combat at $15:56$, then this avatar is regarded active at hour 15 and 16. We used playtime of individual avatars to further calculate playtime overlap between avatar pairs. The playtime overlap was used as the indicator of propinquity (control variable) for predicting collaboration ties between avatars (please refer to the Introduction). Because the intention of controlling for propinquity is to control for *potential* of collaboration between avatars with regard to their active online time, we believe that an hour-level (instead of minute- level) overlap better indicates the potential.[4]

   We made some transformations on the raw values for later analyses: First, for the nonzero playtime overlaps, we took the logarithm of the raw numbers plus one, so that a 1-h overlap would correspond to $\ln(1+1)$ and the largest overlap in the current sample (15-h overlap) to $\ln(15+1)$. This transformation was conducted out of the consideration that while longer overlap might grant more potential for collaboration because, for example, two friends may be more likely to notice one another and have more time to game together if their online times overlap more, this increase should not be linear. Second, we

[2] As players can have more than one avatar, there were 5897 players behind the 6376 avatars. Of these 5897 players, 5452 had only one avatars, 411 had two avatars (153 one male and one female; 181 two male; 77 two female), and 34 had three avatars (13 all male; 2 all female; 19 mixed-gender avatars). Thus, of the players who had more than one avatars, 61% of them had avatars that were of the same gender.

[3] The theoretical basis of snowball sampling is the real-life phenomenon and the theory of small worlds, which states that people are well-connected in social networks and that the steps taken to reach to anyone in a network is often quite small (for reference, the theoretical number of steps needed is six [37]). Assuming that the combat network of *Destiny* is well connected, our 18 sampling-wave process was theoretically adequate to explore the population.

[4] To control for the potential of collaboration between avatars, we want to know if two avatars are active online around the same time so that they have the opportunity to collaborate. We only have information on combat encounters, each of which lasts for a few minutes. People do not engage in combat one immediately after another; there are often time gaps between combats, during which the avatars are online but are doing leisure things. Calculating playtimes at the minute level is too narrow to reveal ones active time in general; instead, taking the hour-level information can better portray when the avatars are active and thus whether avatar pairs have the potential to collaborate. For example, if, according to the combat histories, one avatar is in combat from $15:35$ to $15:45$ and the other avatar from $15:46$ to $15:56$, then these two avatars should have the potential for collaborating, since their active times are similar, but a minute-level overlap would miss this.

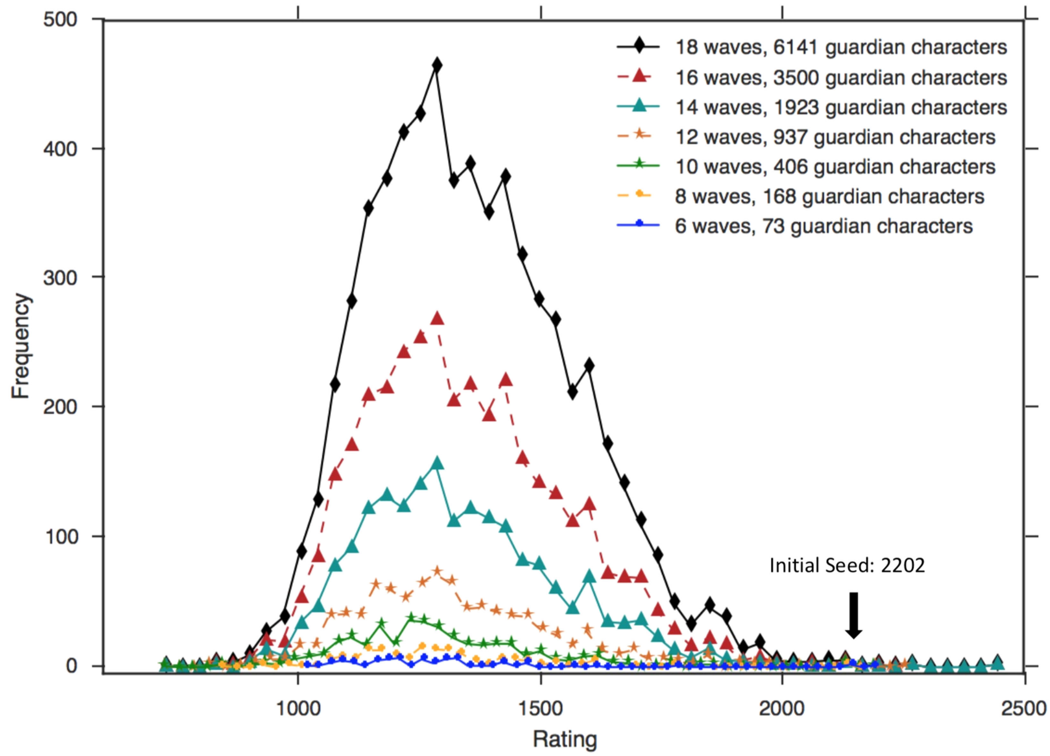


Fig. 3. Distributions of avatar rating as the snowballed sample including more waves. (Avatars who were controlled by the same player had the same rating.)

assigned the value -3 to zero-hour overlap. We chose this value because, on the one hand, we wanted to recognize the fact that no overlap in playtime means no opportunity for collaboration; on the other hand, we avoided an overly small number (e.g., -10), else the assigned value would effectively dichotomize the scale by dwarfing the effect of the nonzero overlaps, since the other end of the scale is $\ln(15+1)$, 2.8. Because of the transformations, we only regarded the playtime overlap as a control variable and did not substantially interpret the estimated effect of this factor.

6) *Combat Rating:* Recall that the ranking on the leaderboard from which we scraped data is based on combat rating. While all the above-mentioned variables were included in the statistical models (explained below), combating rating was not because the rating is at the player level and the analyses are at the avatar level. Despite this, a closer look at the distribution of combat rating helps to reveal how representative the current sample is in terms of players performance as compared to the *Destiny* player population who played ranked PvP. On the leaderboard, the first (rating $=$ 3273) and the last player (rating $=$ 341) were the outliers of the rating distribution. Excluding these two players, other players ratings ranged from 2684 to 551. As can be seen in Fig. 3, although we started from an avatar of a high-rated player, the network quickly crawled downward on the leaderboard and even included the lowest-rated players. The sample seemed to have reached an equilibrium in terms of players performance (i.e., the shape and statistical features of the distribution were stabilized).

C. *Procedure of Analysis*

As mentioned above, we used ERG modeling [58] to test the effect of gender on collaboration networks while controlling for the effects of various confounding network-formation mechanisms. The ERG model is a statistical method that can take the interdependence of network ties into account. The estimation of ERG models is not based on closed equations, as is common in regression, but on a stochastic simulation procedure named Markov chain Monte Carlo maximum likelihood estimation (MCMCMLE) [58]. Using the specified predictors (i.e., mechanisms that are believed to be relevant to the observed network), the procedure tries to replicate the features of the observed network in the networks that it simulates. In other words, the goal of the estimation is to find coefficients based on which the network features of the simulated network distribution are centered on those of the observed network. Following the estimation goal, the goodness of fit of ERG models rests on evaluating how well the simulated networks capture features of the observed networks. More detailed criteria for model evaluation are presented in Section IV along with model results. In Appendix, we provide a short and more statistical description of ERG models based on the introductory book by Lusher *et al.* [58]. For detailed explanations, please refer to the book as well as a special volume in the *Journal of Statistical Software* introducing the `statnet` package in `R` by Handcock *et al.* [64].

The effect of gender on collaborations is the focus of this study. To test the hypothesis of avatar gender homophily on collaborations, two gender terms were estimated: The *gender homophily* effect corresponded directly to hypothesis testing; it modeled whether ties were more likely to exist between avatars that were of the same gender. Meanwhile, to control for the possibility that one gender in general collaborated more than the other gender, a *gender main effect* was also modeled. It is standard to include both main effect and homophily effect if the goal were to investigate a homophily pattern [58].

Three attributes, avatar class, performance, and number of combats played were included to control for characteristic intersection with gender. For these attributes, we again estimated both main effect and homophily effect. We tested whether Hunters or Warlocks were more likely to collaborate than Titans (the main effect of avatar class), whether higher-performed avatars were more likely to collaborate than the lower-performed (the main effect of avatar performance), whether avatars who were in more combats in a day were more likely to collaborate than the ones who were in fewer (the main effect of number of combats); in addition, we also tested whether avatars of the same class (the homophily effect of avatar class), of similar performance levels (the homophily effect of performance), and played similar amount (the homophily effect of number of combats) were more likely to collaborate.

To control for the endogenous networking mechanisms that influenced the distribution of network ties, we included three statistical terms. An *edge* term estimated the baseline propensity of tie occurrence (comparable to the intercept in linear regression). An *alternating star*[5] term represented preferential attachment by statistically modeling the degree distribution of the network; that is, how centralized the network is around the high-degree nodes. An *alternating triangle*[6] term represented the effect of triangle closure.

ERG modeling is computationally expensive and, thus, it is infeasible to fit one ERG model on a 6376-node network. Instead, based on the method suggested by Stivala *et al.* [66], we drew 40 small samples from the data set, fitted ERG models on them separately, and used meta-analyses to combine the results of the small samples. We again used snowball method to construct these small samples, each of which was based on five random initial seeds and three waves of snowball chain. To ensure that all collaboration partners within these three waves were in the data set, the initial seeds were restricted to the first 15 waves (recall that the current 6376-avatar network was constructed through 18 waves of snowballing).

The average size of the 40 small samples is 118 avatars (range: 71–197). For fitting the ERG models on the 40 small samples separately, the predictive terms were specified as the same, but settings controlling the MCMCMLE (i.e., burn-in, interval and sample size) were adjusted for the different sample sizes, so that the autocorrelations between the lags were smaller than 0.4 [58]. All ERG models were fitted in R using the `statnet` package [67].

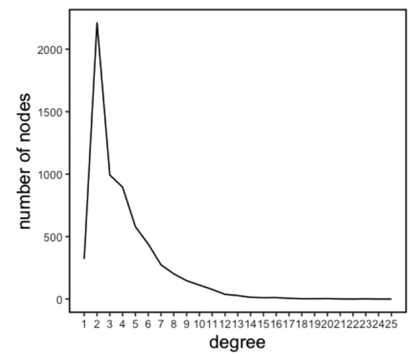


Fig. 4. Distribution of the number of degree of the overall network (6376 avatars.)

To summarize the results from the separate ERG models, we adopted the random-effect model of meta-analyses [68] as previously done by Stivala *et al.* [66]. The meta-analysis was estimated using the R package `metafor` [69].

IV. Results

We first present results of a preliminary analyses on how male and female avatars differed on specific variables. Subsequently, we report the main results of the ERG models via a meta-analysis, and demonstrate the goodness of fit of the ERG models by specifically focusing on an illustrative model.

A. Preliminary Analyses

We conducted a series of tests to examine differences between male and female avatars on their avatar class, performance level, number of combats, and playtime hours. Male and female avatars significantly differed in avatar class, $\chi^2(2) = 28.40$, $p < .001$; the male avatars' percentages in Titan, Hunter, and Warlock were 33%, 39%, and 28%, respectively, whereas female avatars' percentages were 27%, 44%, and 30%. More male than female avatars were created as Titans and the opposite is true for Hunters. As for performance level, standardizing the kills/deaths ratio, female avatars ($M = 0.22$) have higher kills/deaths ratio than male avatars ($M = -0.09$), $t(2959.7) = -10.69$, $p < .001$. But male and female avatars did not significantly differ in the number of combats they played (Male $= 13.55$, Female $= 13.94$, $t(3296.4) = -1.63$, $p = .104$) nor in their active playtime hours (Male $= 4.64$, Female $= 4.69$, $t(3338.6) = -0.66, p = 0.51$) on the day when the current data were collected.

For the overall network that consisted of 6376 avatars, there were, in total, 12 700 ties: 1160 were between two female avatars, 5230 between two male avatars, and 6402 between a female and a male avatar. We performed a chi-square test on the gender composition of collaborative ties without controlling for any covariates, gender was not a factor influencing the formation of collaborative ties at all ($\chi^2(2) = 11.44, p = 0.997$). Fig. 4

[5]We used statnet's term `altkstar(lambda, fixed = TRUE)`. After experimenting different values for the smoothing constant, λ is fixed at 1.2. We refer the reader to [58, p. 71] and [65, p. 8] for a further explanation on smoothing constant.

[6]We used statnet's term `gwesp(alpha, fixed = TRUE)`. After experimenting with different values for the smoothing constant, λ is fixed at 1.3 (i.e., $\alpha = ln(\lambda) = 0.3$). We refer the reader to [58, p. 71] and [65, p. 13] for a further explanation on smoothing constants.

TABLE I
ESTIMATES AND TEST STATISTICS OF THE META-ANALYTICAL SUMMARIES OF THE 38 CONVERGED ERG MODELS

	Estimate	SE	p-value	95%CI(lower)	95%CI(upper)
Edges	7.64	0.70	< .0001	6.26	9.01
Gender(main)	0.01	0.03	0.84	-0.05	0.06
Gender (homophily)	0.10*	0.04	0.02	0.01	0.18
Class (Hunter; main)	-0.03	0.02	0.22	-0.08	0.02
Class (Warlock; main)	0.01	0.03	0.66	-0.04	0.06
Class (homophily)	0.07*	0.03	0.02	0.01	0.12
K/D ratio (main)	0.09***	0.02	< .0001	0.06	0.12
K/D ratio (heterophily[a])	-0.22***	0.03	< .0001	-0.27	-0.17
Number of matches (main)	-0.14***	0.01	< .0001	-0.17	-0.12
Number of matches (heterophily)	0.05**	0.02	0.01	0.01	0.08
Overlap in playtime	1.53***	0.07	< .0001	1.40	1.66
Triangle closure (alternating triangle)	3.28***	0.08	< .0001	3.12	3.43
Preferential attachment (alternating star)	-6.96***	0.28	< .0001	-7.51	-6.41

[a]In ERG models, the homophily effects of continuous variables are included conversely as heterophily effect, so that a negative effect indicates homophily.

presents the distribution of the number of degrees (the number of collaborative ties that each avatar had).

B. Results of Meta-Analysis

To address the research question of how gender, along with other mechanisms, predicted the distribution of the collaboration ties, we estimated 40 ERG models separately and used a meta-analysis to summarize the results. In total, 38 of the ERG models converged.[7] Integrating the results with the random-effect model of meta-analyses, the results are presented in Table I. In ERG models, the homophily effects of continuous variables are included conversely as heterophily effect, so that a negative effect indicates homophily.[8]

The current research question asks how gender influences collaboration relationships, we found a significant gender homophily effect ($B = 0.10$, $SE = 0.04$, $p = .02$). This number indicates that conditioning all other ties of the network, the logit of a tie, i.e.,

$$\log\left(\frac{\text{probability of a tie existing}}{1 - \text{probability of a tie existing}}\right)$$

was increased by 0.1 if the tie was between the same gender. The significant result lent support to our hypothesis that real-life gender homophily could also be found in the virtual environment. There was no significant gender main effect, suggesting male and female avatars formed a similar amount of collaboration ties. We remind the readers that in the preliminary section above, we performed a chi-square test on the gender composition of the ties (comparing the observed versus expected numbers of female–female, male–male, and male–female ties) and the chi-square was nonsignificant ($p = .996$). The difference between the significant result of the ERG models and the nonsignificant results of the chi-square test suggested that controlling for the set of confounders was relevant.

In regard to the control variables, we found that avatars who were of higher performance level ($B = 0.09$, $SE = 0.02$, $p < 0.0001$) and gamed fewer combats ($B = -0.14$, $SE = 0.01$, $p < 0.0001$) had more collaborative ties. In addition, avatars of the same class ($B = 0.07$, $SE = 0.03$, $p = 0.02$), of similar performance levels ($B = -0.22$, $SE = 0.03$, $p < 0.0001$), and played different amount of combats ($B = 0.05$, $SE = 0.02$, $p = 0.01$) were significantly more likely to collaborate with each other. Playtime overlap was a significant factor influencing the probability of collaboration ties ($B = 1.53$, $SE = 0.07$, $p < 0.0001$). But given the transformations we performed on this variable, its coefficient is not easily interpretable. Comparing to the avatar features, endogenous network mechanisms showed much larger effect. Triangle closure was positively significant ($B = 3.28$, $SE = 0.08$, $p < 0.0001$), indicating that avatars had a higher tendency to game with collaborative partners collaborative partners. The preferential attachment was negatively significant ($B = -6.96$, $SE = 0.28$, $p < 0.0001$), suggesting that the ties did not gathered around a small number of popular avatars. This is reasonable given the nature of our network. Typically, highly popular nodes can only form if the marginal effort needed to maintain additional network ties is small (e.g., a popular account on *Instagram*). Yet, our network ties require substantial effort from both end of the relationship and, thus, it is almost impossible to observed high-degree avatars in our network since all activities in our study were observed in one day.

C. ERG Models Goodness of Fit

We present the summary result of the 38 converged models above. In this section, we report the goodness of fit of the ERG

[7]The two models that did not converge were of the size 84 and 196, respectively. In the work of Stivala *et al.* [66] from which we adopted the current method of using a meta-analysis to summarize ERG models, they also reported that not all models converged and proceeded with the ones that did converge.

[8]For a categorical feature (e.g., gender), a positive estimate indicates a homophily effect because the feature is modeled as a binary variable: 1 indicates that the node pairs share the same category (e.g., both female or both male) and 0 otherwise. A positive effect, therefore, means that, compare to different-categories avatar pairs, the same-category pairs are more likely to share ties. For a continuous feature (e.g., performance level), a positive estimate indicates a heterophily effect because the feature is modeled as absolute differences between avatar pairs on that feature (e.g., the absolute difference between avatar pairs on their performance levels). A positive estimate, therefore, indicates that the pairs that differ more on the particular feature are more likely to share ties (i.e., a heterophily effect).

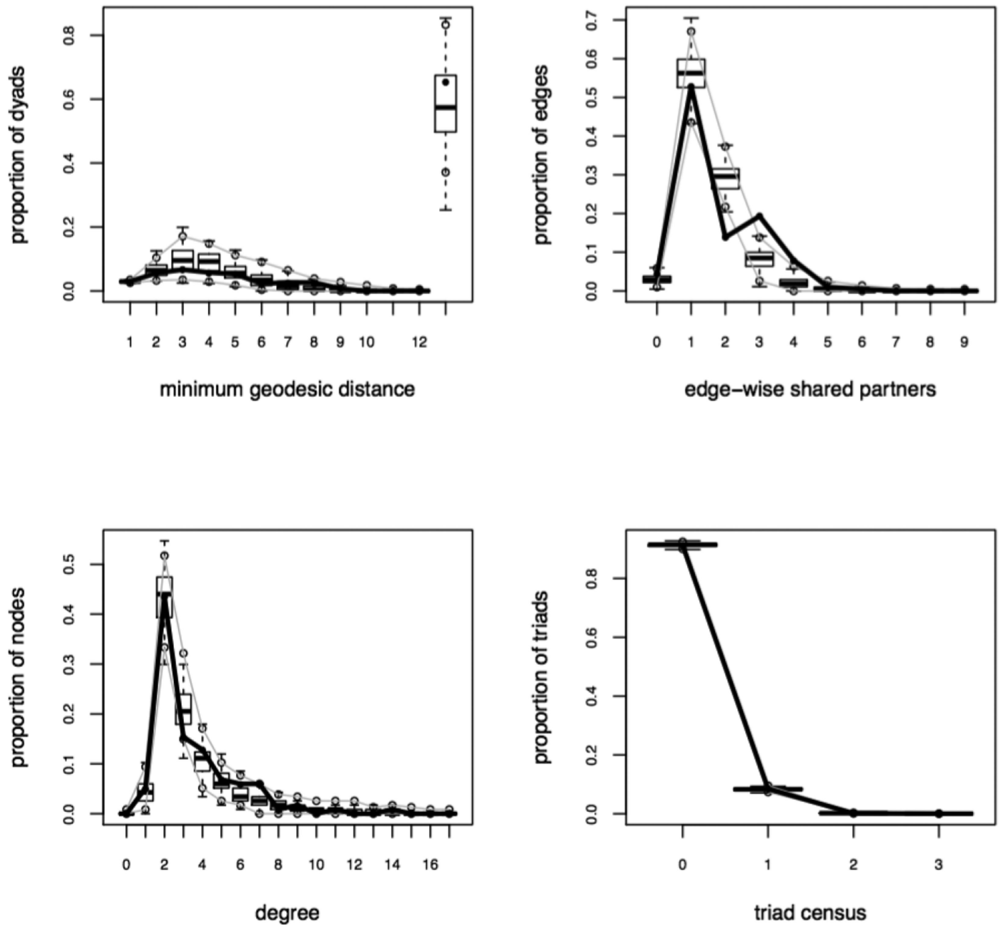


Fig. 5. Graphical summaries of four goodness-of-fit indicators: The geodesic distance distribution, the edgewise shared partner distribution, the degree distribution, and the triad census distribution.

models. Given that we repeated the same modeling method on the 38 samples, we do not report the model fit for all 38 models separately; instead, as an example, we report the goodness of fit of one model, which was based on a sample of 117 avatars. This model is referred to as the "example model." The 38 models that successfully converged have the similar goodness of fit to this example model.

The goal of ERG model fitting is to find coefficients based on which the statistics of the simulated network distribution are centered on those of the observed network. Therefore, one approach to examine the goodness of fit of a model is to simulate a series of networks based on the estimated coefficients and examine how well these simulations replicate the observed network. Using the `gof` function of `statnet`, 100 networks were simulated. For the example model, the goodness of fit of the model statistics showed that the specified features of the observed network were all nonsignificantly different from the means of the corresponding features of the simulated network distribution (p-values range from 0.84 to 1.00). Therefore, the predictive terms specified for the observed model were well captured by the simulated networks of the coefficients. This was the case for all 38 converged models. The modeled features of the 38 observed networks were all nonsignificantly different from the averaged features of the simulated networks.

Moreover, a well-fitted model should capture some important features of the observed network regardless of the model specification. That is, even if some network features are not explicitly modeled, they should still be well represented by a good model. The `gof` function offers goodness of fit assessment for four network features: 1) the geodesic distance distribution—the proportion of pairs of nodes whose shortest connecting path is of length k, k = 1, 2, 3, ...; 2) the edgewise shared partner distribution—the proportion of pairs of nodes who share k partners, k = 1, 2, 3, ...; 3) the degree distribution—the proportion of nodes of k degree, k = 1, 2, 3, ...; and 4) the triad census distribution—the proportion of 3-node sets having k edges among them, k = 0, 1, 2, or 3 [70]. Fig. 5 shows the graphical summaries of these four goodness-of-fit indicators for the example model. The gray lines depict the range of the network features based on the 100 simulated networks and the bold black line depicts the network features of the observed network. In the four plots, the black lines are mostly within the range of simulated statistics. This suggests that, although the model abstraction is not perfect, the simulations based on this model can reproduce important features of the observed network. Although we cannot present the plots for the 38 converged models here separately, observed patterns across all models are highly similar.

V. Discussion

In conclusion, while the two investigated genders by themselves did not differ in their general level of collaboration (i.e., female avatars did not have more collaboration ties than male avatars, and vice versa), we found evidence supporting our hypothesis of gender homophily: Avatars of the same gender were more likely to collaborate with each other (i.e., female avatars collaborate more with female and male avatars collaborate more with male), but it is important to point out that this effect was a weak one (coefficient $= 0.1$, $p < 0.05$).

Given the context of the current finding, this observed tendency of gender homophily is likely to be driven by preferences toward the same-gender others. As we discussed before, in real life, interactions and relationships are gendered not only because people may genuinely prefer the same-gender others, but may also because the same-gender people enjoy similar activities, are socially pressured to befriend each other, take similar (e.g., occupational) positions, and, in turn, are physically closer to each other [7], [14], [33], [37]. Yet, in the game context, these reasons are likely to be weakened; for example, players who voluntarily enter a game should not have incompatible activity interests, and the anonymity of the game may greatly free people's interactions from social pressures [2], [38]. As such, if we consider why there is a gender homophily pattern in MMOGs, the psychological preference could be one motivation.

We want to emphasize that one advantage of the current study is that we controlled for other tie-formation mechanisms [60]. Some other characteristics, which also affect collaboration ties, may intersect with gender, whereby confound the observed gender effect [57], [58]; we controlled for the influences of several important ones (avatar class, performance level and the amount of gameplay). In addition, considering that players whose playtime did not overlap would have no potential for collaboration, we controlled for playtime overlap between avatar pairs. We also took heed of the influence of endogenous networking mechanisms and controlled for the effect of triangle closure and preferential attachment [58], [59]. The results of the preliminary analyses and the effects of the control variables support that modeling the confounding mechanisms was meaningful. This is because there were significant differences between male and female avatars on some controlled characteristics and the control variables had significant effects on the probability of collaboration ties.

In regard to implications of our findings, theoretically, since research has always been interested in disentangling various motivations behind gender homophily [7], [14], the current finding on the basis of a virtual environment may contribute to this pending question by highlighting the relevance of the preference. Practically, since the social element has become a crucial aspect of MMOGs [8], game developers may be interested in the nature of social interactions on these platforms. Inferring on the basis of the current finding, interactions in the virtual environment may mirror some important features of the real-life society.

Besides serving for the gender effect, the findings of the preliminary analysis and control variables also generated interesting insights. We highlight a few. First, we found that male avatars were more likely to be chosen as Titans than females, whereas female avatars were more likely to be Hunters than males. This distinction seems to hold connotations of gender stereotypes because Titans are designed to be strong and carrier of heavy armors and Hunters are agile and good at combat by stealth. While the specialty of Hunters is more ambiguous, Titan's emphasis of physical strength is a stereotypical masculine trait. The observation that more male avatars indeed took Titan role suggests that people did consider what is typically manly and womanly when designing avatars, just like previous research has shown [71].

Second, we found that female avatars had a higher performance level, in the term of kills/deaths ratio, than male avatars. This result seems counter-intuitive, as anecdotal evidence suggests that men are better and more hardcore gamers than women. Yet, our finding does not stand alone. A large-scale survey study on *EverQuest* players has also reported that women played more hours than men and that women contained a larger percentage of hardcore gamers [55]. To reconcile such findings with the contrasting evidence, which showed that MMOGs are male-dominated environments (i.e., the majority is male), [3], [41], [47], [55], one possibility is that the female players are different from females in the general population. Precisely because MMOGs are commonly considered as male activities and social norms tend to discourage females to play games [47], [55], many women whose motivations are not strong enough are filtered out by the higher threshold of entering the game, leaving only the highly motivated ones in games.

In contrast to differences in performance level, in the current sample, we did not find significant differences between males and females with regard to their number of combats and playtime, but this is likely due to our sampling method. Unlike performance level, which reflects an accumulative effort, number of combats and playtime were calculated based on players activities on a single day, so it is imaginable that no gender difference is detected upon one-day activeness.

Third, concerning the effect of avatar class on collaboration ties, the three avatar classes did not differ in the amount of collaboration, but same-class avatars were significantly more likely to collaborate. Though this effect was weak, it is interesting because it differs from the common perception that different classes tend to collaborate since their skills complement one another. One possible explanation is that the current collaboration ties only included the relatively strong relationships (we define collaboration as coplaying in at least three combats on one day); players of the similar mental propensity were more likely to choose the same class, appreciate each other, and, thus, form stronger relationships. In actual combats, together they game with random strangers who play other classes. This is, of course, only a speculation, but if this were true, then it may suggest a link between preference/propensity and relationship formation with avatar class being the medium.

Fourth, we found highly significant effects for triangle closure and preferential attachment. The former indicated that ties in the collaboration networks formed triangles and avatars had a tendency to collaborate with teammates, whereas the latter offered counter-evidence to preferential attachment, meaning that collaboration ties were evenly distributed and did not cluster around a few highly popular avatars. These two findings together suggest that, on the one hand, avatars did collaborate in small groups since their interactions closed triangles but, on the other hand, these relationships did not lead to cliques, since ties were evenly distributed across the network.

A. Limitations and Future Work

There are a few caveats with regard to the current study. The first one concerns the interpretation of avatar gender. Based on the existing observations in the literature, we suggest that the concept of avatar gender inherits meanings from real-life society. Studying how virtual gender associates with network ties might reveal how the concept of gender plays in role in online interactions. However, it is important to recognize the variety in motivations of choosing avatar gender; not every motivation is substantial and some motivations even deliberately go against our suggestion that male and female avatars carry meanings of men and women, respectively. For example, based on our personal experiences and observations, players can choose a certain gender merely because they like the outfit or voice of that specific character; or players may find female avatars fun because it is counter-stereotype to operate her very aggressively. A comparison between how relationship formation is shaped by players actual gender versus their avatars gender might shed light on how peoples interpersonal relationships are constrained by innate characteristics and how they would be different if possible. Unfortunately, we do not have access to players actual gender in this study. Future studies may consider making this interesting comparison. For pragmatic reasons, being that *Destiny* offered only two gender options for the avatar customization and, again, we did not have player information, our investigation treated gender as binary. We acknowledge that this does not fully reflect the complexity of gender in real life. It would be interesting for future research to investigate nonbinary or even ungendered options (e.g., robotic avatars) offered by other games.

Second, the current sample was drawn from the PvP leaderboard and were limited to activities to a single day (Saturday). For one, we chose Saturday since we observed more activities on Saturday but precisely because of that, a Saturday might not be representative of days in general. For another, players on the leaderboard may not be representative of the *Destiny* population. Some players engaged in the PvE part of the game or play more casually and do not partake in ranking. Previous research studying motivations behind gameplay have suggested that female players are more likely to be motivated by social reasons, whereas male players tend to be more achievement-oriented [3], [55]. Therefore, since we sampled from the leader board, which is a competitive context, our sample might be biased to contain less casual players or even less female players specifically. In addition, since competitive gameplay is the goal of PvP, the look of the avatars might not be very important for the PvP players and findings on avatars based on PvP might not be generalizable to the PvE players. Therefore, our target population is carefully defined as the PvP network on Saturday and inference should not be made beyond this population.

Third, when estimating the ERG models, we could not control for the (3-wave) snowball structure of the samples. The chain-structure of snowball data is by nature different from the structure of the population network, since relationships in a population are not chain-sequenced; therefore, controlling for the snowball structure of the sample is important if the goal were to estimate characteristics of population based on the available samples [72]. There have already been studies that published possibilities of controlling for the snowball structure, but we could not find a suitable method in the currently released statistics packages. For example, Handcock and Gile [73] proposed a method to conditionally model the snowball structure by treating the snowballed data equivalent to missing-at-random data. However, this method requires the size of the population graph to be reasonably small and the number of nodes known; neither requirement was satisfied by the current data set. Pattison *et al.* [72] proposed a different method that controls snowball structure by taking this structure into account during model estimation. For example, when simulating networks, ties are not allowed if they skip over waves (e.g., a wave-1 node cannot have a relationship with a wave-3 node). This method has been implemented on empirical data sets (e.g., see [66]). However, we have not yet found this function available in the released version of PNet (software).

Fourth, our data were scraped from a publicly accessible website and are a kind of organic data/found data (in contrast to designed data) that were not originally designed for the purpose of research [74]. Thus, we were confronted with obstacles that are common to organic data in general. For example, organic data often do not have clearly defined variables and users have to design and decide how to exact information from the data [75]. Specific to the current study, our dependent variable collaborative ties between avatars was such a construct that we extracted from the data. We used the number of shared combats between pairs of avatars and assume that coplaying in at least three combats indicates intentional relationships. However, we do not know how *Destiny*'s system worked and it is possible that avatars were indeed randomly assigned to play with the same strangers for more than three times. Reversely, by setting the cutoff point, we missed out the collaborations that were intentional but occurred less than three times. Thus, our dependent variable was subjected to measurement error although we were constructing a complete network for the sample.

Despite the above limitations, the current study has shown that the tendency to form collaboration ties was related to avatar gender. A gender homophily pattern was observed on top of various confounding mechanisms. In a virtual fantasy world where people partook anonymously and interacted through avatars, interpersonal relationships seemed to be gendered analogously to the real-life interactions.

Appendix
Short description of ERG models

ERG models are tie-based models that help to understand the forming mechanisms of a social network. ERG models share some similarities to logistic regression, and comparing it to logistic regression may help to explain the basic principle of ERG models. ERG models analyze tie status (i.e., tie=1/ no tie=0) between pairs of avatars as the dependent variable. They are similar to logistic regression in the sense that ERG models also estimate the probability for the binary dependent variable (i.e., tie existence) given the predictive terms. However, while logistic regression is estimated using closed equations, estimations from ERG models rely on stochastic simulation procedures; correspondingly, ERG models can take fundamentally different predictive terms compared to logistic regression.

Logistic regression assumes independence in the units of the dependent variable and, thus, it can take predictive variables if the units of the predictive variables are also independent from each other. For example, in the current context where the units of dependent variable are the collaboration ties, logistic regression can predict the probability of tie/no-tie using predictive terms such as the gender difference between avatar pairs. This is because the difference in gender for one pair of avatar is independent of the difference in gender for another pair of avatars. While these effects—referred as actor attribute effects (of the actors)—can be modeled by logistic regression, these are often not sufficient for understanding the forming mechanisms of social networks because actor attributes are not the only mechanisms underlying social networks and ties in social networks are not independent of each other (i.e., the assumption of independence in the units of the dependent variable is violated) [58]. For example, the theory of triangle closure suggests that a tie is more likely to exist if it closes triangles in the network. ERG models address the limitation of logistic regression in the network context and take the interdependence between network ties into account by including endogenous networking mechanisms in the model. This is achieved by specifying the logit of tie/no-tie as conditional logit and including statistics of network features as the predictive term

$$\log\left(\frac{P(X_{ij}=1|X_{-ij}=x_{-ij})}{P(X_{ij}=0|X_{-ij}=x_{-ij})}\right) = \theta_1\delta^{+}_{ij,1}(x)+\theta_2\delta^{+}_{ij,2}(x) + \cdots + \theta_p\delta^{+}_{ij,p}(x). \quad (1)$$

X_{ij} refers to the tie between avatar *i* and *j*; X_{-ij} refers to all other ties in the network besides the tie between avatar *i* and *j*. The left-hand side of the equation denotes that the logit of the tie i and j depends on the rest of the network. The right-hand side of the equation specifies how the logit is conditional on the rest of the network: $\delta^{+}_{ij,p}(x)$ denotes *change statistics* for the network feature p. Taking the feature of triangle closure as an example, the statistic of this feature can be operationalized as the number of triangle counts[9]; the change statistics of this feature (i.e., $\delta^{+}_{ij,\text{for triangle closure}}(x)$) is the number of triangles closed by adding a tie between avatar *i* and *j*. If $\theta_{\text{for triangle closure}}$ is positive and significant, then the tie between avatar *i* and *j* will have higher probability to exist if it can close triangles in the network. By modeling change statistics, the probability of ties between avatar pairs are conditioned on the rest of the network.

Equation (1) focuses on the probabilities of individual ties and the predictors are the changes in the network configurations elicited by changes in the individual ties. Since a network is a joint form of all individual ties, (1) can be equivalently written as the probability for a network with predictors being the counts of (instead of changes in) the network features

$$P(X=x)=\frac{1}{\kappa}\exp(\theta_1 z_1+\theta_2 z_2+\cdots+\theta_p z_p). \quad (2)$$

$X=x$ refers to a specific network x. z_p is the count of network feature p. Take triangle closure as an example, $z_{\text{for triangle closure}}$ is the overall count of triangle in network x. κ is the normalizing term, which transforms the right-hand side of (2) to a probability score.

The predictors in right-hand side of (2) specify the effect of network configuration. In ERG models, predictive terms for the effects of actor attributes can be included in the form of interactions between the attribute y and the ties x: $\theta_a z_a(x,y)$. Taking gender as an example of the actor attribute, the main effect of gender can be included in the model by specifying $z_{\text{main effect of gender}} x, y$ as $\Sigma_{i<j} x_{ij}(y_i+y_j)$ (x_{ij} denotes the tie between avatar *i* and *j*; y_i denotes the gender of avatar *i*; y_j denotes the gender of avatar *j*); the gender homophily effect can be included in the model by specifying $z_{\text{gender homophily}} x, y$ as $\Sigma_{i<j} x_{ij} y_i y_j$. The ERG modeling uses MCMCMLE procedure [58] to find coefficients (θ) for the predictive terms. The estimation algorithm first simulates a distribution of networks based on a starting vector of coefficients; it then updates the vector of coefficients until the simulated network distribution is centered on the observed network. In other words, the goal of the model fitting is to find coefficients, based on which the average network statistics of the simulated network distribution are not different from those of the observed network. The estimates, then, give the maximal support to the data observed.

[9]In the analyses, we use alternating triangle to control for triangle closure and alternating star for preferential attachment. The network statistics of these two terms is an alternatingly weighted sum of the count of k-triangle and k-star, respectively. We do not bring up "alternating" here for simplicity.

Acknowledgment

The author(s) received no financial support for the research, authorship, and/or publication of this article. The authors wish to acknowledge the kind feedback of Dr. Marijtje van Duijn and colleagues on the subject of exponential random graph analysis.

References

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

[2] H. Cole and M. D. Griffiths, "Social interactions in massively multiplayer online role-playing gamers," *CyberPsychol. Behav.*, vol. 10, no. 4, pp. 575–583, 2007.

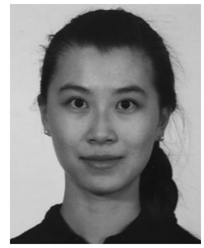
[3] N. Yee, "The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments," *Presence: Teleoperators Virtual Environ.*, vol. 15, no. 3, pp. 309–329, 2006.

[4] N. Yee, "Motivations for play in online games," *CyberPsychol. Behav.*, vol. 9, no. 6, pp. 772–775, 2006.

[5] S. T. Fiske and S. L. Neuberg, "A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation," *Adv. Exp. Soc. Psychol.*, vol. 23, pp. 1–74, 1990.

[6] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, no. 1, pp. 415–444, 2001.
[7] C. M. Mehta and J. Strough, "Sex segregation in friendships and normative contexts across the life span," *Develop. Rev.*, vol. 29, no. 3, pp. 201–220, 2009.
[8] N. Ducheneaut and R. J. Moore, "The social side of gaming: A study of interaction patterns in a massively multiplayer online game," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 2004, pp. 360–369.
[9] P. M. Blau, D. Ruan, and M. Ardelt, "Interpersonal choice and networks in china," *Social Forces*, vol. 69, no. 4, pp. 1037–1062, 1991.
[10] D. C. French, U. Purwono, and A. Triwahyuni, "Friendship and the religiosity of indonesian muslim adolescents," *J. Youth Adolescence*, vol. 40, no. 12, pp. 1623–1633, 2011.
[11] E. E. Maccoby and C. N. Jacklin, "Gender segregation in childhood," *Adv. Child Develop. Behav.*, vol. 20, pp. 239–287, 1987.
[12] M. B. Parten, "Social play among preschool children," *J. Abnormal Social Psychol.*, vol. 28, no. 2, pp. 136–147, 1933.
[13] M. Rhee, D. Yang, and T. Yoo, "National culture and friendship homophily in the multinational workplace," *Asian Bus. Manage.*, vol. 12, no. 3, pp. 299–320, 2013.
[14] C. L. Martin, O. Kornienko, D. R. Schaefer, L. D. Hanish, R. A. Fabes, and P. Goble, "The role of sex of peers and gender-typed activities in young children's peer affiliative networks: A longitudinal analysis of selection and influence," *Child Develop.*, vol. 84, no. 3, pp. 921–937, 2013.
[15] L. Lee, C. Howes, and B. Chamberlain, "Ethnic heterogeneity of social networks and cross-ethnic friendships of elementary school boys and girls," *Merrill-Palmer Quart.*, vol. 53, no. 3, pp. 325–346, 2007.
[16] J. Stehlé, F. Charbonnier, T. Picard, C. Cattuto, and A. Barrat, "Gender homophily from spatial behavior in a primary school: A sociometric study," *Social Netw.*, vol. 35, no. 4, pp. 604–613, 2013.
[17] J. Strough and A. M. Covatto, "Context and age differences in same- and other-gender peer preferences," *Social Develop.*, vol. 11, no. 3, pp. 346–361, 2002.
[18] F. Poulin and S. Pedersen, "Developmental changes in gender composition of friendship networks in adolescent girls and boys," *Develop. Psychol.*, vol. 43, no. 6, pp. 1484–1496, 2007.
[19] W. Shrum, N. H. Cheek Jr, and S. MacD, "Friendship in school: Gender and racial homophily," *Sociol. Educ.*, vol. 61, pp. 227–239, 1988.
[20] C. M. Mehta and J. Strough, "Gender segregation and gender-typing in adolescence," *Sex Roles*, vol. 63, nos. 3/4, pp. 251–263, 2010.
[21] L. DiDonato and J. Strough, "Contextual influences on gender segregation in emerging adulthood," *Sex Roles*, vol. 69, nos. 11/12, pp. 632–643, 2013.
[22] M. Kalmijn, "Sex segregation of friendship networks. individual and structural determinants of having cross-sex friends," *Eur. Sociol. Rev.*, vol. 18, no. 1, pp. 101–117, 2002.
[23] H. M. Reeder, "The effect of gender role orientation on same-and cross-sex friendship formation," *Sex Roles*, vol. 49, no. 3, pp. 143–152, 2003.
[24] M. Monsour, *Women and Men as Friends: Relationships Across the Life Span in the 21st Century*. Mahwah, NJ, USA: Lawrence Erlbaum Associates Inc, 2002.
[25] C. Barone, "Some things never change: Gender segregation in higher education across eight nations and three decades," *Sociol. Educ.*, vol. 84, no. 2, pp. 157–176, 2011.
[26] B. Zafar, "College major choice and the gender gap," *J. Human Resour.*, vol. 48, no. 3, pp. 545–595, 2013.
[27] A. Hegewisch, H. Liepmann, J. Hayes, and H. Hartmann, "Separate and not equal? Gender segregation in the labor market and the gender wage gap," *IWPR Briefing Paper*, vol. 377, pp. 1–16, 2010.
[28] P. England, "Gender inequality in labor markets: The role of motherhood and segregation," *Social Politics: Int. Stud. Gender, State Soc.*, vol. 12, no. 2, pp. 264–288, 2005.
[29] M. Gamba and B. H. Kleiner, "The old boys network today," *Int. J. Sociol. Soc. Policy*, vol. 21, nos. 8/9/10, pp. 101–107, 2001.
[30] S. McDonald, "What's in the old boys network? Accessing social capital in gendered and racialized networks," *Social Netw.*, vol. 33, no. 4, pp. 317–330, 2011.
[31] M. K. Feeney and M. Bernal, "Women in stem networks: Who seeks advice and support from women scientists?" *Scientometrics*, vol. 85, no. 3, pp. 767–790, 2010.
[32] S. Hanson and M. Blake, "Gender and entrepreneurial networks," *Regional Studies*, vol. 43, no. 1, pp. 135–149, 2009.
[33] G. Kossinets and D. J. Watts, "Origins of homophily in an evolving social network," *Amer. J. Sociol.*, vol. 115, no. 2, pp. 405–450, 2009.
[34] M. A. Caldwell and L. A. Peplau, "Sex differences in same-sex friendship," *Sex Roles*, vol. 8, no. 7, pp. 721–732, 1982.
[35] E. E. Maccoby, *The Two Sexes: Growing up Apart, Coming Together*, vol. 4. Cambridge, MA, USA: Harvard Univ. Press, 1998.
[36] E. H. Gorman, "Gender stereotypes, same-gender preferences, and organizational variation in the hiring of women: Evidence from law firms," *Am. Sociol. Rev.*, vol. 70, no. 4, pp. 702–728, 2005.
[37] C. Kadushin, *Understanding Social Networks: Theories, Concepts, and Findings*. Cary, NC, USA: OUP USA, 2012.
[38] J. A. Bargh, K. Y. McKenna, and G. M. Fitzsimons, "Can you see the real me? Activation and expression of the true self on the internet," *J. Social Issues*, vol. 58, no. 1, pp. 33–48, 2002.
[39] C. Beavis and C. Charles, "Would the realgirl gamer please stand up? gender, lan cafés and the reformulation of the girlgamer," *Gender Educ.*, vol. 19, no. 6, pp. 691–705, 2007.
[40] E. Castronova, "Virtual worlds: A first-hand account of market and society on the cyberian frontier," CESifo Working Paper Series, pp. 1–41, Dec. 2001. [Online]. Available: http://ssrn.com/abstract=294828
[41] M. D. Griffiths, M. N. Davies, and D. Chappell, "Breaking the stereotype: The case of online gaming," *CyberPsychol. Behav.*, vol. 6, no. 1, pp. 81–91, 2003.
[42] M. L. Lewis, R. Weber, and N. D. Bowman, "They may be pixels, but they're my pixels: Developing a metric of character attachment in role-playing video games," *CyberPsychol. Behav.*, vol. 11, no. 4, pp. 515–518, 2008.
[43] K. Christoph, H. Dorothée, and V. Peter, "The video game experience as true identification: A theory of enjoyable alterations of players' self-perception," *Commun. Theory*, vol. 19, no. 4, pp. 351–373, 2009.
[44] J.-K. Lou, K. Park, M. Cha, J. Park, C.-L. Lei, and K.-T. Chen, "Gender swapping and user behaviors in online social games," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 827–836.
[45] R. M. Martey, J. Stromer-Galley, J. Banks, J. Wu, and M. Consalvo, "The strategic female: Gender-switching and player behavior in online games," *Inf., Commun. Soc.*, vol. 17, no. 3, pp. 286–300, 2014.
[46] Z. Hussain and M. D. Griffiths, "Gender swapping and socializing in cyberspace: An exploratory study," *CyberPsychol. Behav.*, vol. 11, no. 1, pp. 47–53, 2008.
[47] N. Yee, "Maps of digital desires: Exploring the topography of gender and play in online games," in *Beyond Barbie and Mortal Kombat: New Perspectives on Gender and Gaming*, K. Yasmin, H. Carrie, D. Jill, and S. Jennifer, Eds. Cambridge, MA, USA: MIT, 2008, pp. 156–183.
[48] M. Lehdonvirta, Y. Nagashima, V. Lehdonvirta, and A. Baba, "The stoic male: How avatar gender affects help-seeking behavior in an online game," *Games Culture*, vol. 7, no. 1, pp. 29–47, 2012.
[49] P. R. Messinger, X. Ge, E. Stroulia, K. Lyons, K. Smirnov, and M. Bone, "On the relationship between my avatar and myself," *J. Virtual Worlds Res.*, vol. 1, no. 2, pp. 1–17, 2008.
[50] M. Snyder, E. D. Tanke, and E. Berscheid, "Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes," *J. Personality Social Psychol.*, vol. 35, no. 9, pp. 656–666, 1977.
[51] J.-E. R. Lee, C. I. Nass, and J. N. Bailenson, "Does the mask govern the mind?: Effects of arbitrary gender representation on quantitative task performance in avatar-represented virtual groups," *Cyberpsychol., Behav. Social Netw.*, vol. 17, no. 4, pp. 248–254, 2014.
[52] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore, "Alone together?: Exploring the social dynamics of massively multiplayer online games," in *Proc. SIGCHI Conf. Human Factors Comput. Syst*, 2006, pp. 407–416.
[53] J. Pirker, A. Rattinger, A. Drachen, and R. Sifa, "Analyzing player networks in destiny," *Entertainment Comput.*, vol. 25, pp. 71–83, 2018.
[54] A. Rattinger, G. Wallner, A. Drachen, J. Pirker, and R. Sifa, "Integrating and inspecting combined behavioral profiling and social network models in destiny," in *Proc. Int. Conf. Entertain. Comput.*, 2016, pp. 77–89.
[55] D. Williams, M. Consalvo, S. Caplan, and N. Yee, "Looking for gender: Gender roles and behaviors among online gamers," *J. Commun.*, vol. 59, no. 4, pp. 700–725, 2009.
[56] M. Szell and S. Thurner, "How women organize social networks different from men," *Scientific Reports*, vol. 3, 2013, Art. no. 1214.
[57] A. Wimmer and K. Lewis, "Beyond and below racial homophily: Erg models of a friendship network documented on facebook," *Amer. J. Sociol.*, vol. 116, no. 2, pp. 583–642, 2010.
[58] D. Lusher, J. Koskinen, and G. Robins, *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[59] G. Robins, *Doing Social Network Research: Network-Based Research Design for Social Scientists*. Newbury Park, CA, USA: Sage, 2015.
[60] Y. Huang, C. Shen, and N. S. Contractor, “Distance matters: Exploring proximity and homophily in virtual world networks,” *Decision Support Syst.*, vol. 55, no. 4, pp. 969–977, 2013.
[61] A. L. Jia, S. Shen, R. V. D. Bovenkamp, A. Iosup, F. Kuipers, and D. H. Epema, “Socializing by gaming: Revealing social relationships in multiplayer online games,” *ACM Trans. Knowl. Discovery Data*, vol. 10, no. 2, 2015, Art. no. 11.
[62] G. Stafford, H. Luong, J. Gauch, S. Gauch, and J. Eno, “Social network analysis of virtual worlds,” in *Proc. Int. Conf. Active Media Technol.*, 2012, pp. 411–422.
[63] M. Szell, R. Lambiotte, and S. Thurner, “Multirelational organization of large-scale social networks in an online world,” *Proc. Nat. Acad. Sci.*, vol. 107, no. 31, pp. 13636–13641, 2010.
[64] M. S. Handcock, D. R. Hunter, C. T. Butts, S. M. Goodreau, and M. Morris, “Specification of exponential-family random graph models: Terms and computational aspects,” *J. Stat. Softw.*, vol. 24, no. 4, p. 1548, 2008.
[65] M. Morris, M. S. Handcock, and D. R. Hunter, “Specification of exponential-family random graph models: Terms and computational aspects,” *J. Stat. Softw.*, vol. 24, no. 4, pp. 1548–7660, 2008.
[66] A. D. Stivala, J. H. Koskinen, D. A. Rolls, P. Wang, and G. L. Robins, “Snowball sampling for estimating exponential random graph models for large networks,” *Social Netw.*, vol. 47, pp. 167–188, 2016.
[67] S. M. Goodreau, M. S. Handcock, D. R. Hunter, C. T. Butts, and M. Morris, “A statnet tutorial,” *J. Stat. Softw.*, vol. 24, no. 9, pp. 1–27, 2008.
[68] M. Borenstein, L. V. Hedges, J. Higgins, and H. R. Rothstein, “A basic introduction to fixed-effect and random-effects models for meta-analysis,” *Res. Synthesis Methods*, vol. 1, no. 2, pp. 97–111, 2010.
[69] W. Viechtbauer, “Conducting meta-analyses in r with the metafor package,” *J. Stat. Softw.*, vol. 36, no. 3, pp. 1–48, 2010.
[70] D. R. Hunter, M. S. Handcock, C. T. Butts, S. M. Goodreau, and M. Morris, “ERGM: A package to fit, simulate and diagnose exponential-family models for networks,” *J. Stat. Softw.*, vol. 24, no. 3, 2008, Art. no. nihpa54860.
[71] R. A. Dunn and R. E. Guadagno, “My avatar and me–gender and personality predictors of avatar-self discrepancy,” *Comput. Human Behav.*, vol. 28, no. 1, pp. 97–106, 2012.
[72] P. E. Pattison, G. L. Robins, T. A. Snijders, and P. Wang, “Conditional estimation of exponential random graph models from snowball sampling designs,” *J. Math. Psychol.*, vol. 57, no. 6, pp. 284–296, 2013.
[73] M. S. Handcock and K. J. Gile, “Modeling social networks from sampled data,” *Ann. Appl. Stat.*, vol. 4, no. 1, pp. 5–25, 2010.
[74] R. Groves, “Designed data” and “organic data,” 2001. [Online]. Available: https://www.census.gov/newsroom/blogs/director/2011/05/designed-data-an d-organic-data.html
[75] M. P. Couper, “Is the sky falling? New technology, changing media, and the future of surveys,” *Survey Res. Methods*, vol. 7, no. 3, pp. 145–156, 2013.



Shiyu Zhang received the M.Sc. degree in data science: business and governance from Tilburg University, Tilburg, The Netherlands, in 2017. She is currently working toward the Ph.D. degree in survey methodology with the University of Michigan, Ann Arbor, MI, USA.

She was with the Department of Cognitive Science and Artificial Intelligence, Tilburg University.

Sander Bakkes received his Ph.D. degree from Tilburg University, The Netherlands, in 2010.

He is an Assistant Professor with the Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands. His research interests include personalized and adaptive game experiences.

Diederik Roijers received his Ph.D. degree from the University of Amsterdam, The Netherlands, in 2016.

He is a Senior Lecturer in technical computer science with the HU University of Applied Science Utrecht, Utrecht, The Netherlands, and a Researcher with the AI lab with the Vrije Universiteit Brussel, Brussels, Belgium. His research interests include reinforcement learning and planning, and multiobjective decision making.

Pieter Spronck received his Ph.D. degree from Maastricht University, The Netherlands, in 2005.

He is a Full Professor of computer science with the Department of Cognitive Science and Artificial Intelligence, Tilburg University, Tilburg, The Netherlands. His research interests include game analytics and adaptive artificial intelligence for video games.