Computers in Human Behavior 55 (2016) 605–613

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

What's in a name? Ages and names predict the valence of social interactions in a massive online game

Athanasios V. Kokkinakis [a], Jeff Lin [b], Davin Pavlas [b], Alex R. Wade [a, *]

[a] University of York, Psychology Dept, York, United Kingdom
[b] Riot Games, Santa Monica, California, United States

ARTICLE INFO

Article history:
Received 2 June 2015
Received in revised form
20 September 2015
Accepted 22 September 2015
Available online xxx

Keywords:
Video games
Data mining
Age
Personality

ABSTRACT

Multi-player online battle arena games (MOBAs) are large virtual environments requiring complex problem-solving and social interaction. We asked whether these games generate psychologically interesting data about the players themselves. Specifically, we asked whether user names, which are chosen by players outside of the game itself, predicted in-game behaviour. To examine this, we analysed a large anonymized dataset from a popular MOBA ('League of Legends') – by some measures the most popular game in the world.

We find that user names contain two pieces of information that correlate with in-game social behaviour. Both player age (estimated from numerical sequences within name) and the presence of highly anti-social words are correlated with the valences of player/player interactions within the game.

Our findings suggest that players' real-world characteristics influence behaviour and interpersonal interactions within online games. Anonymized statistics derived from such games may therefore be a valuable tool for studying psychological traits across global populations.

1. Introduction

Online video games are played by hundreds of millions of people worldwide and fine-grained statistics on each game are constantly relayed to centralized servers where they can be stored and analysed. These games often require complex team strategies and permit direct personal interactions mediated by real-time chat, as well as inter-player rating mechanisms. They therefore represent a rich potential source of data for psychological investigation.

Previous research on relating personality traits to video game characteristics have often correlated findings from personality questionnaires with game data: either statistics collected within the game environment, or statistics about the amounts or types of games played (Chory & Goodboy, 2011; King, Delfabbro, & Griffiths, 2013; Park, Song, & Teng, 2011; Teng, 2008; Worth & Book, 2014; Yee, Ducheneaut, Nelson, & Likarish, 2011). This approach is valuable because personality questionnaires provide verified indicators about stable, real-life personality traits. However, respondents may respond untruthfully even to questionnaires administered anonymously across the internet and completing these questionnaires is time-consuming, thereby limiting the number of individuals who can be included in each study.

An alternative approach to the psychological analysis of gaming data is to 'mine' very large datasets for scientifically relevant relationships. This approach is interesting for several reasons. First, it is valuable to ask whether large datasets of this type are useful for statistical analysis at all. It may be, for example, that all players adopt a single 'optimal' strategy that leaves little room for personal variability, rendering these datasets uninteresting from a psychological viewpoint.

Secondly, if players do seem to exhibit systematic differences in behaviour, it might be that some of this variance is linked to real-world characteristics such as age, gender or personality (Worth & Book, 2014). Understanding these relationships could provide valuable information about these characteristics at a population level, and this information could be used as a preliminary screen to identify subjects who may be suitable for further testing. Finally, from a system design point of view, if reliable metrics on player behaviour can be established, they can be used to improve the social environment within the game.

1.1. Hypotheses

Here we examined correlations between the valence of in-game

* Corresponding author.
E-mail address: wade@wadelab.net (A.R. Wade).

http://dx.doi.org/10.1016/j.chb.2015.09.034
0747-5632/

interactions and estimates of player age and anti-social tendencies in the massive online battle arena game ‘League of Legends’ (LoL). Here we define anti-social tendencies as being a propensity to engage in behaviour that breaches societal norms and which is likely to cause offense to a large proportion of people.

Because we are harvesting a large, anonymized dataset this study is correlational: We use two pieces of data extracted from usernames and use them to make estimates about the players real-world attributes. We then describe how these estimated variables correlate with in-game behaviour as assessed by the game-based reporting system. We discuss methodological issues relating to the accuracy of these inferences in detail at the end of the paper.

In LoL (Fig. 1) players join small, competing ‘teams’ that proceed to challenge each other for territory and an objective (overtaking the enemy base) in a relatively short time period (typically < 1 h). The precise details of the game are beyond the scope of this paper but there are abundant descriptions in online sources (“League of Legends,” 2013). LoL is currently one of the most popular video games on the planet with an estimated 27 million online players every day (Gaudiosi, 2012). There are regular professional LoL tournaments with prizes worth millions of dollars and top players are eligible for US “internationally recognized athletes” visa status (Blake, 2013).

LoL players communicate through a real-time chat facility. This facilitates coordinated game play but it also allows players to interact socially. Players are also encouraged to evaluate their teammates at the end of each game. For example, players can praise each other for their teamwork or friendliness by sending ‘Honor’. Alternatively they can submit ‘Reports’ chastising other players for deliberately playing badly or sending abusive messages through the chat system. This report system allows us to gather information about the average valence of each player's interpersonal interaction within the game environment. We hypothesized that if players' real-world personality types predict their behaviour within the game, the valence of these interactions might correlate with factors that are related to real-world behaviour. Two such factors are players' ages and their tendency to use foul or offensive language in their public usernames (DeWall, Buffardi, Bonser, & Keith Campbell, 2011; Holtzman, Vazire, & Mehl, 2010) – their ‘anti-social naming tendency’ (ANT).

We analysed players' self-chosen user names to estimate both age and ANT. Many of these user names contained information that informed us about these parameters. Specifically, players often embed their birth date in their user names (e.g. goodplayer1996) and in a separate analysis we show that these dates are highly-correlated with the self-reported ages of the players in the registration procedure. In addition, many usernames contain explicit or lightly obfuscated expletives, racial slurs and boasts that are clearly designed to attract attention (e.g.‘g0ats3x’). Players must invest some time in generating these ANT names as multi-player online games typically have simple filters in place to block straightforward examples of offensive language.

Once we had identified user names that appeared to contain either age or ANT information, we asked if there was a relationship between ANT or age and the average valence of reports that each player sent or received within the game. We found that both age and ANT are predictive of in-game interaction valences as measured by honors and reports. Importantly, we find this effect for both incoming and outgoing ratings (in other words, ratings generated by a player and directed towards other teammates or, alternatively, ratings generated by teammates directed to a player).

2. Methods and materials

2.1. Data sources

Data were provided by the US-based company Riot Games (Santa Monica, CA)—the creators of League of Legends. To improve internet connectivity, Riot Games maintains servers around the world dedicated to particular geographic regions. The data described here were obtained from servers based in North America (NA), Western Europe (EUW), North Eastern Europe (EUNE), Turkey (TK), and Brazil (BR). Riot Games supplied a representative, random sample of 450,000 datasets—one for each player. This large dataset comprised of 100,000 players on each of the NA, EUW, EUNE and BR servers and 50,000 players from the TK servers. The data represent a snapshot of the accounts on the different servers on June 13, 2013. All accounts in the dataset had been created after November 1st, 2012. The number of datasets was chosen to be as large as possible while still remaining computationally tractable.

Our analysis of anti-social user names was based on data from just the NA server (allowing us to identify English language epithets). Our age analysis was based on all available datasets.

Strict controls were imposed of the type of data that were analysed. Data were collected and analysed in accordance with guidelines from both the Association of Internet Researchers (Markham & Buchanan, 2012) and the American Psychological Association (Kraut et al., 2004). It is important to note that only anonymized datasets were analysed. Researchers had no access to personal identifying information and no modification of players' online experience was performed as a result of this research. All players had agreed to Riot's Terms and Conditions as part of the LoL registration procedure and these explicitly allow LoL to use their data for research purposes. All procedures described in this paper

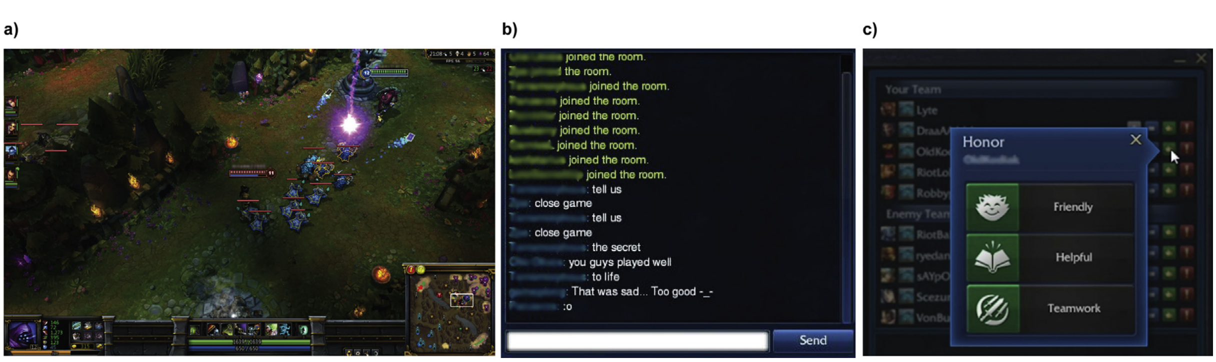

Fig. 1. Game play within League of Legends. a) A screenshot of a LoL match in progress. A small portion of the playing arena is shown (illustrated in the small inset box, bottom right). Individual player ‘summoner names’ or ‘usernames’ appear above the human-controlled characters along with a health indicator. b) In-game chat. Players are able to communicate with each other both during and after a game. c) Sending negative and positive reports is possible after each game ends. Here, a player is choosing to send a positive ‘Honor’ report about a teammate.