Is *Big Brother* Equitable? Diving into Disparities on the Hit Reality Show

Ryan Mitchell

2023-12-16

1 Introduction

Based on George Orwell's dystopian novel *1984*, *Big Brother* is a hit reality television competition that sees a group of contestants enter a house with 24/7 audio and video surveillance. Originally created by John de Mol and launched in the Netherlands in 1999, the show followed twelve players, known as "houseguests", on a 106-day journey where the viewers would periodically vote to "evict" one of them, eliminating them from the competition; when the series was not airing, audiences could view the houseguests' every move via an online livestream. The show came to the United States a year later with the same format, but disappointing ratings for *Big Brother 1* led to a massive shift in approach. From *Big Brother 2* to the most recent installation *Big Brother 25*, the show has given the power to the houseguests, forcing them to evict each other and adding a power struggle component to the fascinating social experiment.

Over the years, *Big Brother* has repeatedly made headlines for problematic behaviors exhibited by both the cast members and production. In 2013, racist and homophobic remarks made by several houseguests on *Big Brother 15* made national news, leading to contestants losing employment and facing extensive backlash from the public. Additionally, the show has historically been scrutinized for a lack of diversity in casting, underscored by the fact that the program produced zero Black winners in its first 22 seasons. Most recently, *Big Brother 25* frustrated audiences as female houseguests won a measly 6 out of 31 weekly competitions; production was also accused of not making proper accommodations for their first deaf houseguest. Fortunately, CBS, the network on which *Big Brother* airs, has taken steps to counteract this history of inequity, namely by launching an initiative to make all reality show casts at least 50% Black, Indigenous, People of Color.

This project will examine the potential influence of certain demographic factors on how a houseguest fares on *Big Brother*, aiming to discover if the program has been equitable for past contestants. Two metrics of performance will be investigated: how a houseguest places on a season (1st, 2nd, 3rd, etc.), and the number of competitions a houseguest wins. How the relationships of demographic factors with performance change over time will also be studied as a means of measuring both the effectiveness of network initiatives and the power of a constantly evolving society.

2 Methodology

2.1 Data

The bulk of the data for this project comes from a dataset created by Vince Dixon as a part of their 2019 data visualization project on diversity in *Big Brother*. Dixon gathered both demographic and performance data on contestants from the first 23 seasons; since there had been no updates made since early 2022, I manually entered the same data for houseguests from *Big Brother 24* and *Big Brother 25*, using the Big Brother Wiki for reference. The resultant dataset contains 357 rows, one for each contestant from the show's second to its twenty-fifth season (houseguests from *Big Brother 1* are omitted due to the show's different format, involving audience evictions and no weekly power competitions.) While Dixon's original dataset contained 49 variables, this was reduced to 17 through both the removal of unneeded columns and the creation of new ones. Data wrangling and cleaning was conducted in both Google Sheets and R.

Variables of note for this project include the following (where a hashtag indicates the variable was not in the original Dixon dataset):

- `age`: The age of the houseguest at time of filming
- `gender`: The gender of the houseguest at time of filming. Note that *Big Brother* has never cast an individual who is openly neither male nor female, but contestants have come out as non-binary after the show.
- `BIPOC`[#]: Binary indicator on whether or not the houseguest is Black, Indigenous, and/or a Person of Color. An "N" in this column refers to white, Anglo-American contestants.
- `LGBT`[#]: Binary indicator on whether or not the houseguest identified as part of the LGBT community at time of filming.
- `placement_bucket`[#]: Ordinal variable indicating how well a houseguest placed on their season. The five levels are "Early Out", "Lower Middle", "Upper Middle", "Late Game", and "Winner"; these each encompass one quarter of the houseguests on each season, except for the fourth quarter which is split into "Late Game" and "Winner".
- `wins_bucket`[#]: Ordinal variable indicating how many competitions a houseguest won on their season, relative to the other houseguests. If multiple houseguests won the same number of competitions, those who were evicted earlier will receive a better rank as they had less opportunities to win competitions. The four levels are "Bottom Quartile", "Low Quartile", "High Quartile", and "Top Quartile"; these each encompass one quarter of the houseguests on each season.

The response variables for this project will be `placement_bucket` and `wins_bucket`. The decision to define a houseguest's placement and competition performance as relative to the other contestants on their season is to account for the fact that different seasons have different amounts of houseguests and competitions; winning 2 competitions on a 10-week season is not identical to winning 2 competitions on a 14-week season.

2.2 Exploratory Data Analysis

To visualize potential disparities in performance on *Big Brother* across demographic groups, I conducted exploratory data analysis
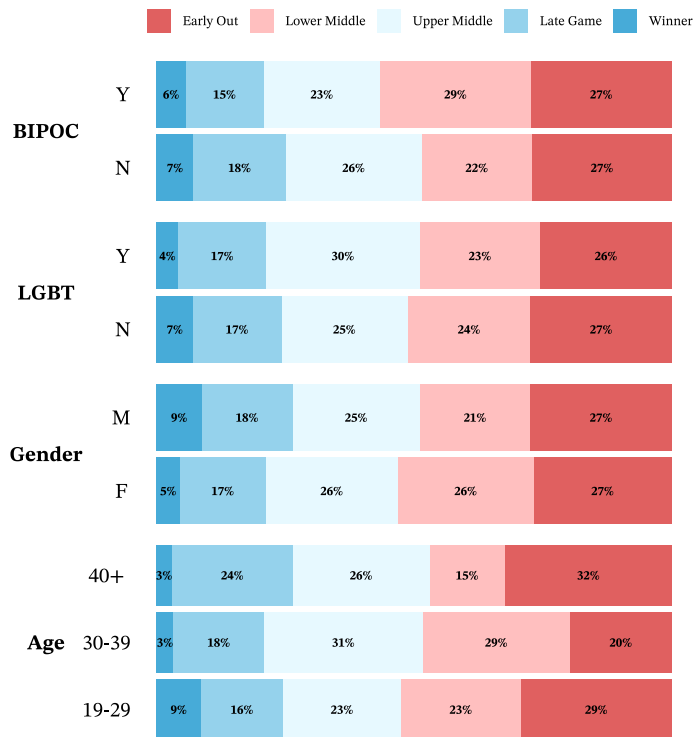
on the 357 houseguests.

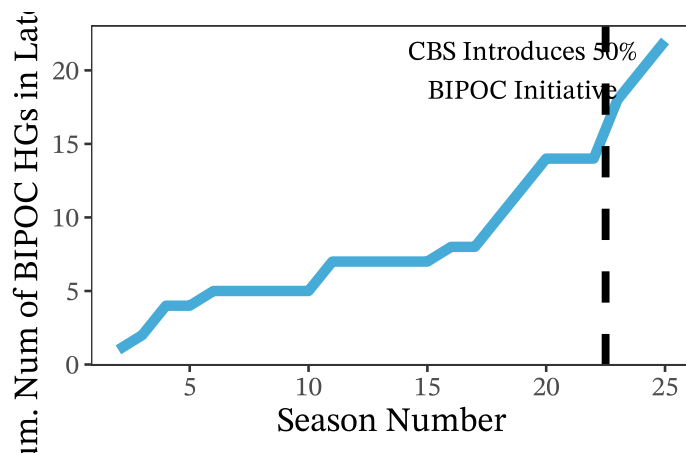Figure 1: Breakdown of placement by demographic group

Figure 2: Cumulative total of BIPOC houseguests making it to the late-game stage of *Big Brother* by season number. Late-game stage is defined as placing in the top 25% of contestants on your season, including the winner.

Figure 1 displays the breakdown of *Big Brother* placement by BIPOC status, LGBT status, gender, and age group; in a scenario with perfect equity, all bars (except Late Game and Winner, which would need to be added together) would be at 25%.