# Title: Data-Driven Player Ratings for Recreational Ultimate Frisbee Leagues

Author: Shane Bussmann

Nearly 5 million people in the United States participate in the sport of ultimate frisbee. Of these, 1.5 million are core participants, meaning that they play ultimate at least 12 times annually. This is more than lacrosse (0.8 million) and rugby (0.4 million) combined. The vast majority of core participants play in recreational leagues in which teams are formed by randomly assigning players to teams such that league-wide parity is maximized.

An integral component of ensuring parity is unbiased player ratings. Historically, player ratings in recreational leagues are based on each individual's self-assessment. Such ratings tend to be biased in two ways. First, experienced players are aware that if they under-rate themselves, they are likely to have stronger teammates and therefore likely to win more games. Second, newer players often over-rate themselves because they are unfamiliar with the league-wide average.

These two biases limit the effectiveness of self-assessment ratings. It is a common occurrence for recreational leagues to have a superstar team that wins every game easily and a team of duds that by halftime of every game is already hopelessly behind. This scenario reduces player retention and limits the growth of the sport of ultimate.

In this contribution, we provide a framework for building data-driven player ratings that are significantly less biased than the self-assessment ratings commonly in place today. The new ratings use the fact that most recreational league players also participate in club leagues in which there is a vetting process that determines whether or not a player can join a team. Simply put, someone who plays on a high-level club team will have a high rating, while someone who plays on a low-level club team will have a low rating.

We apply our methodology to seven years of data from the Boston Ultimate Disc Alliance, the premier ultimate frisbee organization in the greater Boston area. We show evidence that self-assessment ratings are weakly anti-correlated with team performance (as measured by per-game plus/minus score differential; Figure 1). In contrast, ratings based on club experience are strongly correlated with team performance (Figure 2).

These two findings demonstrate that self-assessment ratings should be used as a last resort when no other data are available. We provide guidelines for enacting our data-driven rating system. We envision that the SSAC research paper competition will provide a platform for elevated visibility of this work that will lead recreational ultimate frisbee leagues around the country to adopt our guidelines. Finally, our methodology is generally applicable to any sport in which players participate in a mix of recreational and club leagues.

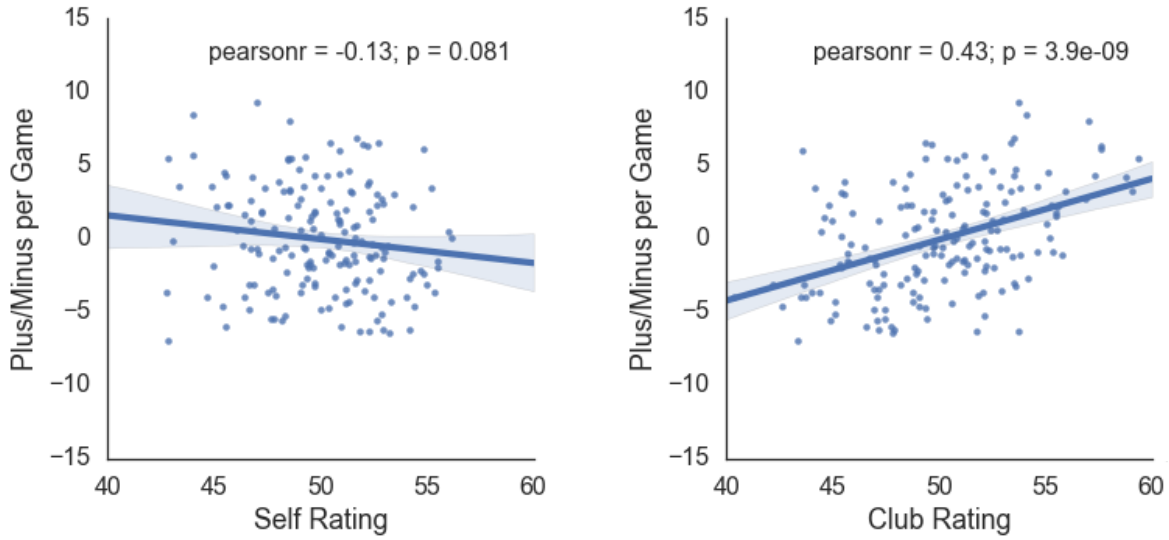


Figure 1: *Left*: Seven years of data from BUDA recreational leagues show that ratings based on self-assessment weakly anti-correlate with per-game plus/minus score differential. *Right*: Ratings based on club experience correlate strongly with per-game plus/minus differential.

500 words