**Analyzing how Athletes and Non-Athletes**

**Responded to the Covid-19 Pandemic**

Josh Garzaniti

*Motivation :*

The Dalai Lama was once quoted saying, “Take care of your mind. It’s where your life comes from.” As of 2024, *Mental Health in America* found that “nearly 60 million Americans had experienced a mental illness over the course of 2020-2021.” (MHIA) The Covid-19 pandemic was a notoriously impactful period for both the mental and physical health of individuals all around the globe. According to a survey done by *Boston College* researchers, by November of 2020, reports of general anxiety in American adults increased up to 50% and depression up to 44%.” (Hayward) Factor analysis on what contributes to increased/decreased mental health attributes is not a novel idea, even more so in the wake of a global outbreak. In fact, the data for this analysis itself originates from a study done on adults in the U.K and Ireland during the pandemic, which focused on comparative analytics between athletes and non-athletes. That said, with the hindsight of factor analysis being used to predict and visualize mental health within the general population, doing the same to analyze how athlete and non-athlete respondents coped with mental health during the height of Covid-19, is a meaningful topic to continue exploring. Continuing to study how adults with different lifestyles handle stressful situations, especially adult athletes, becomes even more meaningful to young adult athletes (usually in college) with the added information that the NCAA in 2022 put out a statement saying, “The number of student-athletes in 2021-22 jumped to over 520,000 (Which was an all-time high).” (NCAA)

Problem Framing

The problem being addressed in this unique situation is that of mental health and how can we better predict, cluster, visualize and analyze what factors impact an individual’s ability or inability to develop resilience to stress and cope given extraordinary circumstances. Uniquely to this research project is how do athletes vary from non-athletes in terms of their mental health assessments under duration and kinds of relationships exist between both groups when it comes to mutual factors/variables. By clustering athletes into groups, we can visualize them based on a variety of both normative measures and sport-related measures (mostly time spent training and competing). Likewise, visualizing clusters within non-athletes’ population samples in combination with factor analysis helps to gain insight on what actions individuals who don’t participate in sports can do to boost their own resiliency.

As pertains to this study, there are four behavioral health metrics included as derived variables that will be key targets for analyzing: *The Adult Mental Health Continuum–Short Form* (Keyes), *The Hospital Anxiety and Depression Scale* (Zigmond, Snaith), *The Brief Resilience Scale* (Smith et al.), and lastly *The Short Loneliness Scale* (De Jong Gierveld, Van Tilburg). Each of these is a dedicated measure of mental health that can be used as outcome variables for assessment. The MHC-SF correlates to an individuals’ assessed overall mental health performance on a scale of 0-70 where higher numbers indicate a sounder holistic well-being. The HADS measures the severity of potential anxiety and depression symptoms where lower scores up to 7 indicate a healthier normality and values above are usually indicative of either mild or severe symptoms (depending on the range). The BRS test displays an individual’s perceived ability to rebound from challenges in life (Again higher score here are better). Then lastly, the SLS displays how isolated an individual perceives they are (This is on a 0-6 range where 0 represents “Least Lonely” and 6 represents “Most Lonely”).

With these measures in mind, and the challenges being “How to map metal health in athletes/non athletes” alongside “How to better identify, predict, and prescribe key factors that correlate with an individual’s ability to stay mentally healthy”, it is only fitting that a key goal here should be to try to maximize scores on the MHC-SF and BRS, while also minimizing scores on the HADS and SLS evaluations respectively. Since these scores are indicative of an individual’s mental state, it is more than worthwhile to explore patterns and trends within that data collected to see if it can lead to possible optimization in health and human habits/practices going forward.

Initial Data Overview

To start off the data collected for this initial assessment was produced as part of a study done by a group of four scholars out of the U.K, which was published back in May of 2021. The data as a result of survey efforts was collected from June 23rd, 2020−July 23th, 2020 (14–16 weeks from the start of lockdown, and notably 2 weeks after restrictions were lifted for those shielding in their homes in the U.K). Data was pulled from the SurveyMonkey application, and the survey itself was given out to a total of 753 adult participants (Knowles, Shannon, Prentice, Breslin). Of those 753 participants, 186 came from Ireland, while the other 558 came from the U.K. Note that as a part of the original data frame, nine individuals noted that they lived outside of either of those two territories, and as such were negated for the final analysis originally produced but are still present within the starting data frame.

The remaining sample is made up of 360 athletes, and 384 non-athletes (which will make for a good respondent group to target analysis on how mental health impacted both groups). It’s also noted that this same sample is made up of 347 male respondents and 397 female respondents (again another good split). Within the original data itself there are a few glaring issues. One of which are the respondents from different countries outside of GB and the UK. For prior research purposes, those were dropped, which might also happen for this study. Another main issue is the formatting of the original data. For starters, there are quite a few statistical outliers in key columns such as the behavioral health metrics. These differ from entries denoted with a “999” because it appears that the “999” entries in the original data frame were intended to represent either 0’s or N/A’s (depending on which variable they’re in). The statistical outliers were handled by first using a function to protect columns that didn’t have egregious maxes within their respective summaries and then sweeping over the remaining columns to remove any observation in which the entry was more than 2 standard deviations away from the mean of each respective numerical variable.

After that, the “999” s were simply mutated to represent either N/A’s if the variable was a survey question unrelated to time spent or a clear binary outcome (in those cases, they were mutated to 0’s). Another part of the cleaning process dealt with entries in the “MentalHealth” column (most likely an indicator of mental health conditions or episodes as factors). Most entries had a numeric range of 2-8, but there were multiple entries that consisted of multiple levels/factors separated by a “,”. So, to address these, a string split was used to break up values and create binary numeric columns representing each Mental health factor and whether it was attributed to a respondent. Then tables were created to initially analyze subsample breakdowns of the population. Sport Level was then mutated to transform back into a clearer ordered level version. Initial correlation tests using the *biserial.cor* function (since Athletes is a binary column and mental health measurements are continuous data types) show that there are weak positive correlations between being an athlete and having stronger: mental health overall scores, stronger resiliency scores, lower loneliness scores, but also higher anxiety scores). Please note all of these were found to not be significant via a cor test.

Bibliography

De Jong Gierveld, J., and T. Van Tilburg. *De Jong Gierveld Loneliness Scale--Short Version (DJGLS)*. 2006. APA PsycTests.

Hayword, Ed. "Anxiety and Stress Spike During Pandemic." *Boston College News*, 2025, <https://www.bc.edu/bc-web/bcnews/campus-community/faculty/anxiety-and-stress-spike-during-pandemic.html>.

Keyes, Corey L. M. *Mental Health Continuum--Short Form (MHC-SF)*. 2009. APA PsycTests.

Knowles C, Shannon S, Prentice G and Breslin G (2021) Comparing Mental Health of Athletes and Non-athletes as They Emerge From a COVID-19 Pandemic Lockdown. *Front. Sports Act. Living* 3:612532. doi: 10.3389/fspor.2021.612532

Mental Health America. "State of Mental Health in America." *Mental Health America*, 2025, <https://mhanational.org/issues/state-mental-health-america>.

NCAA. "NCAA Student-Athletes Surpass 520,000, Set New Record." *NCAA*, 5 Dec. 2022, <https://www.ncaa.org/news/2022/12/5/media-center-ncaa-student-athletes-surpass-520-000-set-new-record.aspx>.

Smith, B. W., et al. *Brief Resilience Scale*. 2008. APA PsycTests.

Zigmond, A. S., and R. P. Snaith. *Hospital Anxiety and Depression Scale (HADS)*. 1983. APA PsycTests.