

Sports Psychology, Film, and the Analysis of Baseball Data.

Arlie R. Belliveau

Email: arlie@yorku.ca

History & Theory of Psychology, York University

ABSTRACT

This paper considers the impact of the first two instances of sports psychology on professional American baseball. What quantifiable changes can be observed after initiatives such as scientific managers Frank and Lillian Gilbreth's 1913 micromotion filming of the New York Giants, or Coleman Griffith's 1938 filming and study of the Chicago Cubs? Repeated measures MANOVA and mixed design analyses revealed no differences between pitcher strikeouts and batter hits before, during, and after the psychologist interventions. Visualization of the data through GoogleVis helps explain the trajectory of this data over time.

Keywords: Sports psychology, MANOVA, baseball, data visualization, Frank and Lillian Gilbreth, Coleman Griffith

1 INTRODUCTION

Sports psychology became division 47 of the APA in 1986, but historians argue that it has held the same general form since the early 1960s (Green & Benjamin, 2009). Professional athletes looking to gain a psychological advantage have been known to hire sports psychologists; and, although this was not common practice until the 1960s, psychology has been applied to sport since 1880 (Green, 2003; Bäuml, 1997 cited in Green & Benjamin, 2009). In baseball, psychologists have offered their expertise concerning reaction times, pitching speeds and styles, proper motions, practice regimens, and overcoming mental barriers since as early as 1913 (*New York Tribune*, June 1913 cited in Belliveau, 2011).

Green (2003, 2009) has argued, however, that the earliest individual psychologists who applied their craft to sport cannot rightfully be considered sports psychologists in the same way that we would consider the term today. He argues that their work was not sustained after their deaths, and more importantly, that their efforts did not yield notable results. This paper considers that claim from a statistical perspective by conducting multivariate analyses on the first two instances of psychological intervention on professional American baseball. The goal is to see if these initial interventions did indeed have any effects on pitcher and batter performance.

1.1 Background History

The first psychologists known to have intervened with a baseball team were industrial management consultants Frank and Dr. Lillian Gilbreth. In an effort to sell their industrial psychology methods to a wider audience, they filmed the pitching speed and accuracy, hitting, base stealing, and catching of the Brown University baseball team in April of 1913 (Brown, 2005; Lancaster, 2004). The *Sunday Providence Journal* picked up the story, and it caught the eye of Mr. George H. Daley, sports editor of the *New York Tribune* newspaper. Daley contacted the Gilbreths to see if they would be interested in conducting similar studies of the New York Giants baseball team. And,

on May 31st of 1913 “in the presence of some 20,000 fans, and the ‘tearing’ music of a 70 member Cuban marine band whose battleship was docked offshore” Gilbreth recorded the Giants’ pitching and batting speeds and motions, as well as the time needed to steal second base (F.B. Gilbreth Diaries, April 22, 1913 cited in Belliveau, 2011, p. 91).

Time and motion measurements were reported in the *New York Tribune* newspaper, among others that picked up the story (for example: *Sporting Life*, 1913). It appears, however, that this was the only communication that occurred between the psychologists and the team. Their efforts were not taken seriously, and they were not hired by the Giants to perform further micromotion analyses (*New York Tribune*, June 1913 cited in Belliveau, 2011).

The next instance of sports psychology in American baseball did not occur until 1938 when Mr. Phillip K. Wrigley, owner of both the gum company and the Chicago Cubs baseball team, hired Dr. Coleman Griffith. Griffith had opened the first sports psychology laboratory at the University of Illinois in the 1920s. He had been developing his methods with football and other college teams for several years before losing his funding during the depression (Green, 2003, 2009).

Accepting the contract, Griffith took what was considered then to be high definition films of the players. He wrote extensive reports on their reaction times, and suggested altering practice efforts to mimic the real stress of a game. Unfortunately, due to skeptical players, and an uncooperative team manager, few of his suggestions were implemented, and he eventually lost the contract in 1940 (Green, 2009). Is it possible that the Cubs and Giants were wrong to dismiss these psychologists so soon?

2 METHODOLOGY

2.1 Research Question

The primary research question of this paper asks whether early interventions on professional baseball teams by psychologists had any effect on player statistics. Specifically, I ask if pitcher strike outs (SO) and batter hits (H) improved for the New York Giants (NYG) and Chicago Cubs (CC) from the year prior to the intervention (1912 and 1937 respectively), to the intervention year (1913 and 1938), and the following year (1914 and 1939).

Since sports psychology has become a widely used approach in baseball (Green & Benjamin, 2009), I hypothesize that the teams were wrong not to hire these psychologists. Player statistics should have risen after their interventions, and I expect to find significantly larger increases in the pitching and hitting scores of the treatment groups than in a control team. Furthermore, I expect to find greater gains for the Chicago Cubs than with the New York Giants, due to their extended time with Griffith.

2.2 Evaluating the Data

I obtained the data for this study through the Sean Lahman Baseball Database (available through the Lahman Library in R; Friendly, 2011). It covers all team and player statistics for each American major league team going back as far as 1871 (Lahman, 2012). After obtaining the data, I selected out the pitcher strikeouts (SO) and batter hits (H) for my desired teams in the years prior to, during, and after the psychological interventions. I also chose a control team (the 1912 Boston Red Sox; BRS) based on their similar standing in the pre-intervention year. The BRS were in 1st place, the CC were in second place, and the NYG were in third place in the year prior to intervention (1912 Major League Baseball Season Summary, 2011; 1937 Major League Baseball Season Summary, 2011).

Before starting the primary analysis I checked to ensure the data met the necessary assumptions of the repeated measures multivariate analysis of variance (RM MANOVA) and mixed models designs. I first tested the shape of each variable to ensure normal distribution. Boxplots revealed positively skewed data both for strikeouts and hits, and so I transformed the data using a square root function. Figure 1.1 shows before and after bar graphs for the transformed data at each time period for pitcher strikeouts. Figure 1.2 shows similar graphs for batter hits. The log transformation left the data negatively skewed. I felt that would be misrepresentative, despite the still abnormally distributed square root data.

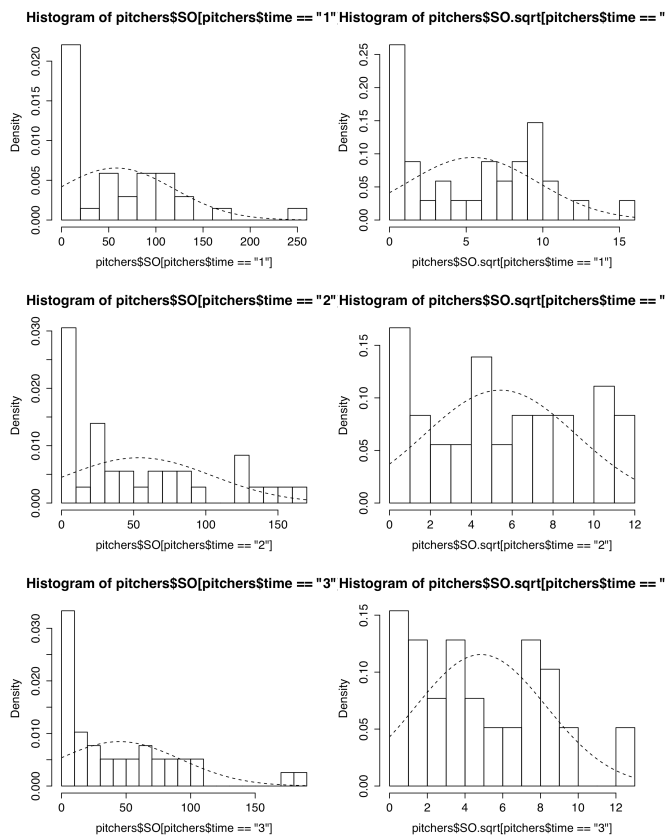


Figure 1.1 Showing the before and after square root transformations for pitcher strikeouts, collapsing over the team variable

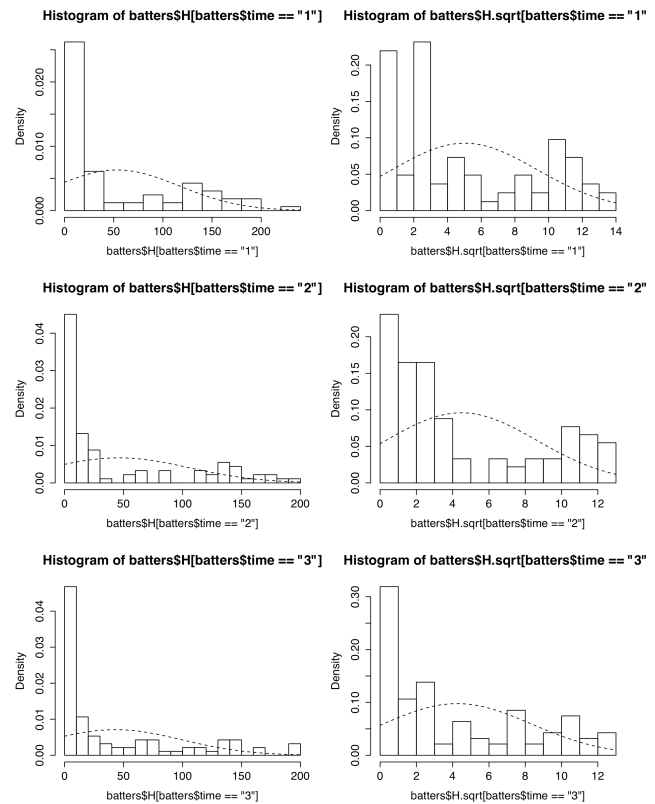


Figure 1.2 Showing the before and after square root transformations for batter hits, collapsing over the team variable

Finally, I tested the homogeneity of the variance covariance matrix for compound symmetry. It was acceptable using the square root transformation, and quite promising for the 'Hits' data.

Sqrt SO	Pre Test	Post Test	Follow-up
Pre Test	17.844		
Post Test	10.402	13.812	
Follow-up	4.717	8.091	11.946

Sqrt H	Pre Test	Post Test	Follow-up
Pre Test	18.586		
Post Test	14.148	17.410	
Follow-up	14.360	14.239	16.669

Due to the historical nature of my research question I was not able to randomize the variables, or run a true experiment. The analysis is further limited by the fact that I used only three teams, and so there are perhaps too few observations to ensure sufficient power.

2.3 Model Selection

RM MANOVA is used when one is looking to compare longitudinal data for multiple groups. It treats the levels of within-subjects independent variables as separate dependent variables (here the pre, post, and follow-up measures). It has the added bonus over a two-way ANOVA of considering the shared variance of the variables, and not requiring the "often-violated assumption of sphericity" (Tabachnick & Fidell, 2007, p. 249). Removing the partial correlations between the three time periods makes the analysis more powerful and increases my chances of

detecting minute differences.

Unfortunately, however, the RM MANOVA is not able to deal with missing data, and it excludes any cases that are not complete. I could have found the missing data for players who were traded, however, that data would come with additional (and undesirable) variability attributable to training received from their new team. Therefore, mixed models designs were run. This strategy has all of the features of the repeated measures model, with the added bonus of utilizing all of the data (Friendly class notes, 2012). Considering the frequency with which baseball players are traded, this method seemed appropriate.

3 RESULTS

The analysis did not yield any significant results, and so I did not follow up with univariate contrasts to test specific hypotheses.

3.1 Repeated Measures MANOVA

A Repeated Measures MANOVA was performed on 13 complete pitcher strikeout score observations. This analysis compared the variables team (NYG, CC, and BRS), time (Pre, Post, and Follow-up) and the team*time interaction. None of these comparisons were significant. For team (collapsing over time), $F(1, 11)=0.8866$, $p=0.367$. For time (collapsing over team) $F(2, 10)=0.0195$, $p=0.981$. The interaction comparing each team over time was also not significant, $F(2, 10)=0.1267$, $p=0.882$. Figure 2.1 shows HE plots for the data.

```
Type III Repeated Measures MANOVA Tests: Pillai test statistic
          Df test stat approx F num Df den Df Pr(>F)
(Intercept) 1 0.70489 26.2737 1 11 0.0003305 ***
team         1 0.07459 0.8866 1 11 0.3666185
time         1 0.00389 0.0195 2 10 0.9807033
team:time    1 0.02472 0.1267 2 10 0.8823478
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

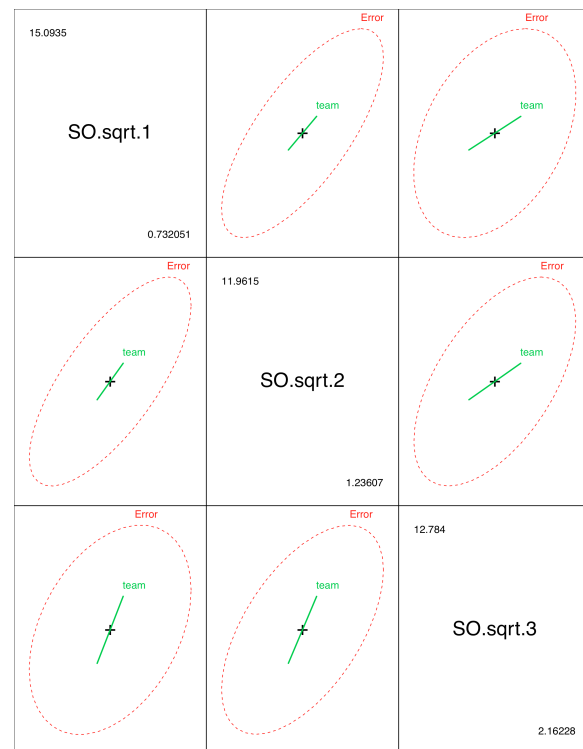


Figure 2.1 Showing HE plots for pitcher strikeouts for each time period

It would appear that pitcher strikeouts actually decreased over the three time periods. The NYG saw a small increase during the treatment year, but lost any gains by the follow-up year. Figure 2.2 demonstrates this below.

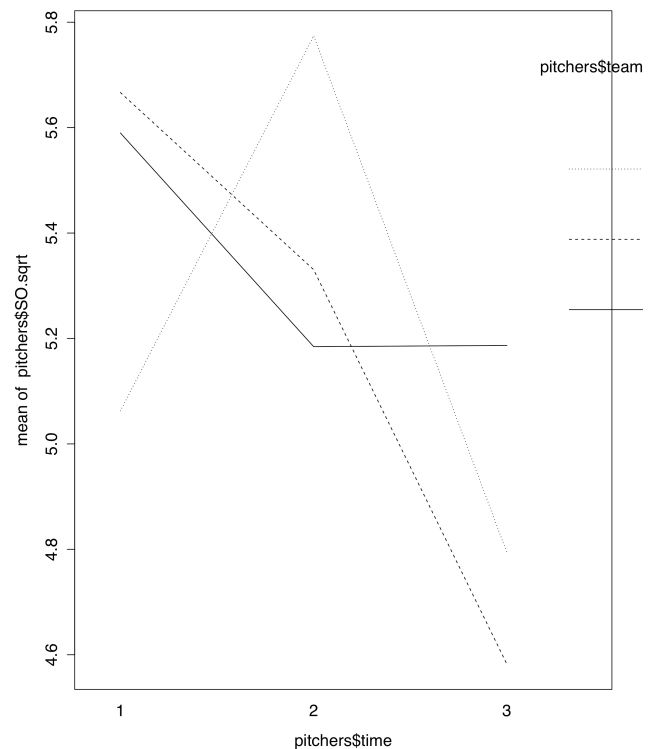


Figure 2.2 shows decreasing strikeouts over the three time periods.

The RM MANOVA on 149 batter hits observations was also not significant. The team comparison was not significant, $F(1,37)=0.6538$, $p=0.4239$. The time comparison was not significant, $F(2,36)=0.3153$, $p=0.7316$. Finally, the interaction comparing each team's hits over time was also not significant, $F(2,36)=0.0751$, $p=0.9278$. Figure 2.3 shows an HE plot for the data.

```
Type III Repeated Measures MANOVA Tests: Pillai test statistic
Df test stat approx F num Df den Df Pr(>F)
(Intercept) 1 0.34978 19.9042 1 37 7.342e-05 ***
team 1 0.01736 0.6538 1 37 0.4239
time 1 0.01721 0.3153 2 36 0.7316
team:time 1 0.00416 0.0751 2 36 0.9278
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

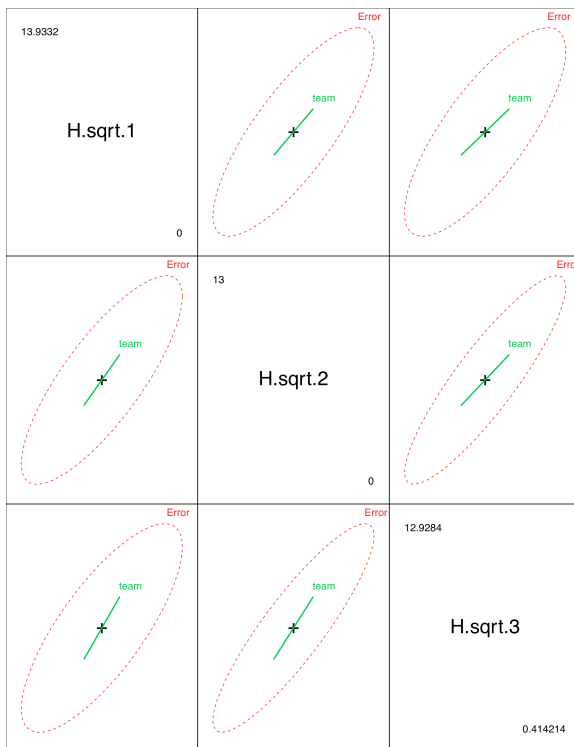


Figure 2.3 Showing HE plots for batter hits for each time period.

As with the pitching data, player hit rates were falling for all three teams over the three time observations. Figure 2.4 uses a line graph to show the general downward trend for each team over the three time periods.

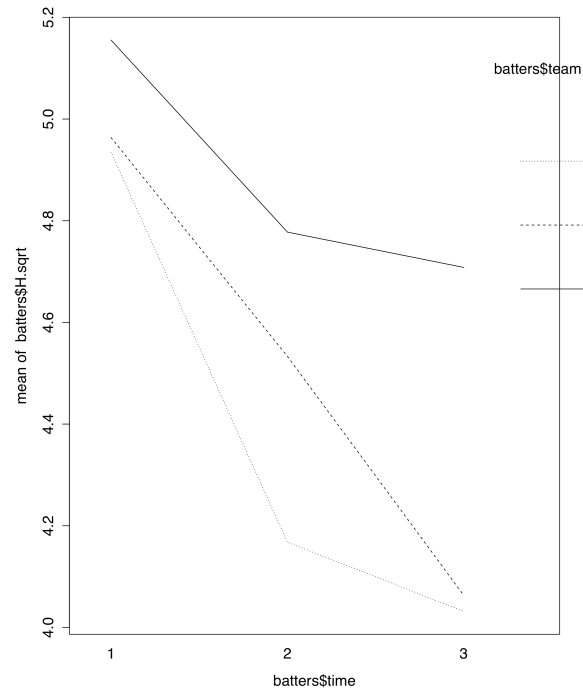


Figure 2.4 Showing a line graph for batter hits over each time period.

3.2 Mixed Design

A mixed models design was conducted on 67 complete and incomplete pitcher strikeout observations. This increased the number of observations by 54. It tested a random slopes and random intercepts model. All of the Wald tests yielded non-significant results. The team variable was not significant, $t(39)=0.842$, $p=0.405$. The time variable was not significant, $t(39)=0.103$, $p=0.912$. And, the team*time interaction was also not significant, $t(39)=-0.171$, $p=0.865$. AIC was 561.7796.

```
Fixed effects: S0.sqrt ~ I(time - 1) * team
              Value Std.Error DF   t-value p-value
(Intercept)  3.0510332 1.6050332 66  1.9009159  0.0617
I(time - 1)   0.0859844 0.8359760 39  0.1028551  0.9186
team          0.6329648 0.7520844 39  0.8416140  0.4051
I(time - 1):team -0.0661403 0.3874341 39 -0.1707137  0.8653
```

Figure 3.1 shows spaghetti plots for each team over the three time periods with the extra data points added in. There is no visible trend.

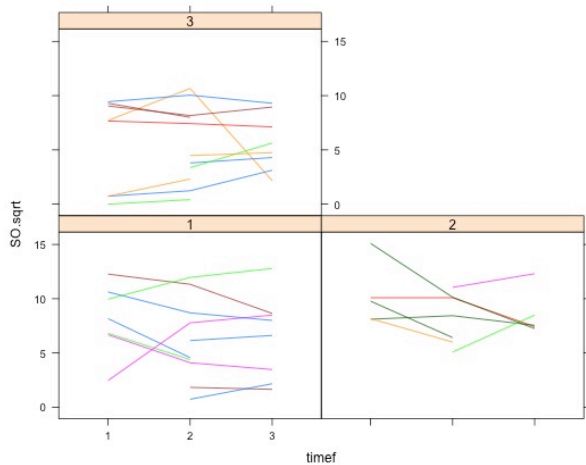


Figure 3.1 shows spaghetti plots of strikeouts for the three teams. Team 1 is the New York Giants, team 2 is the Boston Red Sox, and team 3 is the Chicago Cubs.

To increase power, and test for a more parsimonious model fit, comparisons were conducted between the random intercepts random slopes model (reported above) and models with a fixed intercept (AIC=567.704) and a fixed slope (AIC=606.936). The log likelihood ratio test for both of these tests was significant ($p < 0.01$) indicating that both random intercepts and random slopes were necessary. More parsimonious models could not be used to increase power.

Next, a mixed model design was run on 267 batter hits observations, raising the number of observations by 118. The random slopes random intercepts model did not yield significant results. The t value for the number of hits for each team was not significant, $t(115)=1.173$, $p=0.243$. The time variable was also not significant, $t(115)=-0.089$, $p=0.929$. Finally, the interaction showing team differences over time was not significant, $t(115)=-0.198$, $p=0.844$. AIC=1406.41.

```
Fixed effects: H.sqrt ~ I(time - 1) * team
              Value Std.Error DF   t-value p-value
(Intercept)  2.9105661 1.0180831 148  2.8588689  0.0049
I(time - 1)  -0.0462322 0.5192721 115 -0.0890327  0.9292
team         0.5586802 0.4761103 115  1.1734258  0.2430
I(time - 1):team -0.0482049 0.2437138 115 -0.1977931  0.8436
```

Figure 3.2 displays spaghetti plots for the three teams over each time period. Despite including the missing data values, they do not show any visible trend over time.

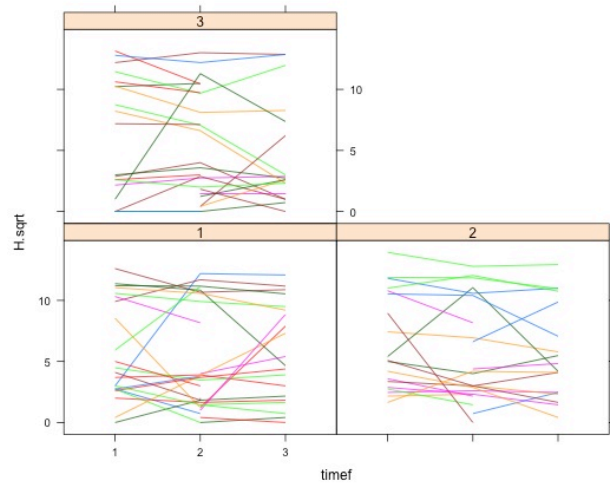


Figure 3.2 shows spaghetti plots of hits for the three teams. Team 1 is the New York Giants, team 2 is the Boston Red Sox, and team 3 is the Chicago Cubs.

Fit comparisons were conducted between the random intercepts random slopes model and models with a fixed intercept (AIC=1404.489) and fixed slope (AIC=1520.73). The log likelihood ratio test for the fixed slope model was significant ($p < 0.0001$) indicating that random slopes were necessary; however, the log likelihood ratio for the random intercept model was not significant ($p=0.3537$).

The final model used random slopes and fixed intercepts, which increased the fit of the model. That increase, unfortunately, was still not enough to find a significant result. The Wald value for time was not significant, $t(115)=-0.906$, $p=0.367$. The test for team was not significant, $t(115)=0.357$, $p=0.722$. And the interaction was not significant, $t(115)=0.377$, $p=0.707$. AIC=1520.73

```
Fixed effects: H.sqrt ~ I(time - 1) * team
              Value Std.Error DF   t-value p-value
(Intercept)  4.636580 0.9633293 148  4.813079  0.0000
I(time - 1)  -0.750657 0.8285326 115 -0.906008  0.3668
team         0.160124 0.4489024 115  0.356701  0.7220
I(time - 1):team 0.146683 0.3886825 115  0.377384  0.7066
```

4 DISCUSSION

In his news story of the micromotion filming of the New York Giants, George Daley joked that if each player became the subject of a micromotion study, managers would have only to look at their instruction cards to determine which rookies would become a “whirlwind” (*New York Tribune*, June, 1913 cited in Belliveau, 2011). Perhaps, had the Gilbreths been given the opportunity, the dazzling player statistics that we see in baseball today would have been made available much sooner.

As it happened - and in answer to my earlier question -, early psychological intervention on baseball did not seem to have a significant effect on pitcher strikeouts or batter hits. So, I cannot reasonably argue to include the Gilbreths or Griffith as proper members of what we would consider

today to be sports psychology.

There are several possible reasons why I found these results. At first I believed that the initial high ranking of the teams may have fallen prey to a regression to the mean. The Chicago Cubs began their pre-treatment season in third place, while in 1912 the New York Giants were in second place, and the Boston Red Sox were in first place. There is a greater distance moving downward in the ranks than there is moving upward.

Another possibility is that the abnormally distributed data impacted the results. Amended player statistics (dividing hits by at bats, or strikeouts by games played, for example) may have helped create more normally shaped data. It is also likely that there was an insufficient number of teams or players to be able to detect the treatment. A thorough power analysis would be quite useful to any future studies.

It could very well be that there really was no treatment effect. The Gilbreths had grossly overestimated pitch speed to be 132 mph. Researchers at the Remington Arms firing range had estimated pitcher Walter Johnson's pitching speed to be 82 mph in 1912 (Thomas, 1995); and so, all of the Gilbreth efforts can be brought into question. The same cannot be said for Griffith though. Perhaps there was something else at play.

After plotting the data in GoogleVis, I was able to see that most teams experienced the same general decline of baseball statistics during my two time periods. It is possible that a confounding variable was impacting these baseball statistics.

The most obvious possibility lies in the fact that the two time periods coincided closely to both World Wars. WWI lasted between 1914-1918, while WWII lasted between 1939-1945. It is possible that the tension building in the years prior to these wars impacted the overall trajectory of baseball statistics. During Griffith's intervention with the Chicago Cubs, the team did rise to make the World Series in 1938 (moving from third to second place). Perhaps it was the onset of the War that prompted their speedy decline in the ranks, overshadowing any positive effects made by Griffith. It would be interesting to see if these teams had a larger proportion of players with familial links to Europe whose abilities may have been impacted by the Wars.

On a final note, other future research might consider the work of later sports psychologists from the 1960s to uncover precisely when their interventions began to significantly influence baseball outcomes. A doubly multivariate design that considers both a measure of strikeouts and hits for the three teams might be attempted.

5 REFERENCES

- [1] Bäumler, G. (1997). The dawn of sport psychology in Europe, 1880-1930: Early pioneers of a new branch of applied science. In (Eds.) Green, C.D., & Benjamin, L.T. (2009). *Psychology Gets in the Game. Sport, Mind, and Behavior, 1880-1960*. Lincoln, N.E.: University of Nebraska Press.
- [2] Brown, E. (2005). *The Corporate Eye: Photography and the Rationalization of American Commercial Culture, 1884-1929*. Baltimore, MA: Johns Hopkins University Press.
- [3] Friendly, M. (2011). Lahman: Sean Lahman's Baseball Database. R Package version 1.0-8. <http://CRAN.R-project.org/package=Lahman>.
- [4] Friendly, M. (2012). Class Notes. PSY6140. York University.
- [5] Gilbreth, F.B. (1885-1924). *Diaries*, Purdue University Libraries. Gilbreth Library of Management Collection. N File, Series 4, Box 3, Folder 1
- [6] Gilbreth F.B., & Gilbreth, L.M., (1912-1924). *Gilbreth Odds and Ends #2*. (Motion Picture). Purdue University Libraries. Gilbreth Collection.
- [7] Green, C.D. (2003). Psychology strikes out: Coleman R. Griffith and the Chicago Cubs. *History of Psychology*. 6(3), 267-283. Doi: 10.1037/1093-4510.6.3.267
- [8] Green, C.D., & Benjamin Jr., L.T. (2009). *Psychology Gets in the Game. Sport, Mind, and Behavior, 1880-1960*. Introduction. The origins of sports psychology. Green & Benjamin (Eds.). University of Nebraska Press.
- [9] Green, C.D. (2009). Coleman Roberts Griffith: "Father" of North American Sports Psychology. p. 202-229. In Green, C.D., & Benjamin Jr., L.T. (2009). *Psychology Gets in the Game. Sport, Mind, and Behavior, 1880-1960*. Green & Benjamin (Eds.). University of Nebraska Press.
- [10] *New York Tribune*. (1913, June). Movies to help baseball players economize force, Purdue University Libraries. Gilbreth Library of Management Collection. Box 5, Folder 2.
- [11] Lahman, S. (2012). *Sean Lahman Baseball Database*. Available: <http://seanlahman.com/>
- [12] Lancaster, J. (2004). *Making Time: Lillian Moller Gilbreth, A Life Beyond "Cheaper by the Dozen"*. Boston, M.A.: Northeastern University Press.
- [13] Perkins, J.S., Gilbreth, L.M., & Barnes, R.M. (1968/2006). *The Original Films of Frank B. Gilbreth*. (Digital Motion Picture). Available at http://www.archive.org/details/OriginalFilm_2 Chicago Chapter of the Society for the Advancement of Management.
- [14] *Sporting Life*. (1913, June 7). 61(14), 23, Purdue University Libraries Gilbreth Library of Management Collection: Box5, Folder1
- [15] *Stock footage – World Series match between New York Yankees and Chicago Cubs in New York*. (1932). (Motion Picture). http://www.criticalpast.com/video/65675026490_baseball-World-Series-New-York-Yankees_Chicago-Cubs_Lou-Gehrig
- [16] *Stock Footage – The New York Yankees defeat the Chicago Cubs 8-3 to win the World Series in New York, United States*. (1938). (Motion Picture). http://www.criticalpast.com/video/65675058999_New-York-Yankees_Chicago-Cubs_baseball-game_winning-world-series
- [17] *Stock Footage – Chicago Cubs warm up at their training camp at Catalina Island, California*. (1938). (Motion Picture). http://www.criticalpast.com/video/65675039943_Chicago-Cubs_baseball-players_players-warm-up
- [18] Tabachnick, B.G., & Fidell, L. S. (2007). *Using Multivariate Statistics*, (5th Ed.). Boston, M.A.: Pearson Education. Allyn and Bacon Publishing.
- [19] Thomas, H.W. (1995). *Walter Johnson. Baseball's Big Train*. Washington, D.C.: Phenom Press.
- [20] 1912 Major League Baseball Season Summary (2011). Sports Reference LLC. Available: <http://www.baseball-reference.com/leagues/MLB/1912.shtml>
- [21] 1937 Major League Baseball Season Summary (2011). Sports Reference LLC. Available: <http://www.baseball-reference.com/leagues/MLB/1937.shtml>

[1] Bäumler, G. (1997). The dawn of sport psychology in Europe, 1880-1930: Early pioneers of a new branch of applied science. In (Eds.)