



4-2013

## American Beisbol: How Cultural Differences Help Explain Different Approaches to Game Playing

Derek Jackson  
*Western Michigan University*

Follow this and additional works at: [https://scholarworks.wmich.edu/masters\\_theses](https://scholarworks.wmich.edu/masters_theses)



Part of the Social and Cultural Anthropology Commons, and the Sports Studies Commons

---

### Recommended Citation

Jackson, Derek, "American Beisbol: How Cultural Differences Help Explain Different Approaches to Game Playing" (2013). *Masters Theses*. 126.

[https://scholarworks.wmich.edu/masters\\_theses/126](https://scholarworks.wmich.edu/masters_theses/126)

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact [wmu-scholarworks@wmich.edu](mailto:wmu-scholarworks@wmich.edu).



AMERICAN BEISBOL: HOW CULTURAL DIFFERENCES  
HELP EXPLAIN DIFFERENT APPROACHES  
TO GAME PLAYING

by

Derek Jackson

A thesis submitted to the Graduate College  
in partial fulfillment of the requirements  
for the degree of Master of Arts  
Anthropology  
Western Michigan University  
April 2013

Thesis Committee:

Robert Anemone, Ph.D., Chair  
Linda Borish, Ph.D.  
Jon Holtman, Ph.D.

# AMERICAN BEISBOL: HOW CULTURAL DIFFERENCES HELP EXPLAIN DIFFERENT APPROACHES TO GAME PLAYING

Derek Jackson, M.A.

Western Michigan University, 2013

The purpose of this thesis project is to examine the effect of culturally derived game strategies on the success level of players in the game of baseball. Specifically, I look at both the influence of how various Latin American cultures teach the game in order to better ensure success of players at the MLB level versus how the game is taught in the United States and Japan. In this way I develop a feedback model in which these game strategies perpetuate a cycle of enculturation that further reinforces cultural/ethnic identities. In order to accomplish this goal I look at the factors that led Latinos to adopt baseball in their culture and how game strategies have been adapted to best get noticed by MLB scouts. In this study, I use advanced statistics referred to as sabermetrics to evaluate player impact, performance, and playtime. These statistics are then used to explore a number of variables such as ethnicity, nationality and age in order to come up with a multi-factorial analysis of the effects of culture on player success.

Copyright by  
Derek Jackson  
2013

## ACKNOWLEDGMENTS

First, I would like to thank some important resources in the completion of this project. While I have never personally met any of the following, I would never have finished without these assets. To the contributors and collaborators at Baseball Prospectus and Royals Review, I thank you.

I would also like to thank the members of my committee, Dr. Robert Anemone, Dr. Linda Borish, and Dr. Jon Holtzman for their guidance, knowledge, and patience. Their support has turned this idea into a true project. I would especially like to thank Dr. Anemone, the chair of my committee, whom I have worked with as a Graduate Assistant and who has been a mentor to me. I would also like to thank Dr. Charles Hilton who was a very important resource in the beginning of this project.

Thirdly, I would like to thank my colleagues, David W. Lee, Bradley P. Aldridge, and Glenn D. Strickland for their assistance in my college career. Your encouragement and assistance has been instrumental in the completion of this project.

Lastly, I would like to thank my future wife for cheering me on and keeping me motivated and on task. Cara, you have made everything much easier.

Derek Jackson

## TABLE OF CONTENTS

ACKNOWLEDGMENTS .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES.....	vi
INTRODUCTION.....	1
BRIEF HISTORY OF MAJOR LEAGUE BASEBALL .....	3
LATIN AMERICAN BASEBALL .....	5
Cuba .....	5
Dominican Republic .....	7
United States .....	9
THE QUESTION OF RACE AND ETHNICITY .....	11
Operationalizing Race and Ethnicity .....	13
CULTURE AND BASEBALL .....	15
Inequalities .....	16
HYPOTHESES .....	18
METHODOLOGY.....	21
RESULTS .....	23
CONCLUSION .....	32

## Table of Contents—continued

### APPENDICES

A. Player Count by Place of Birth.....	34
B. Plot of Means .....	36
C. Tukey Post-Hoc Test for 2B .....	41
D. Tukey Post-Hoc Test for HR.....	43
E. Tukey Post-Hoc Test for GDP .....	45
F. Tukey Post-Hoc Test for BB.....	47
G. Tukey Post-Hoc Test for SO .....	49
H. Tukey Post-Hoc Test for SB .....	51
BIBLIOGRAPHY .....	53

## LIST OF TABLES

1. ANOVA for 2B, using adjusted SS for tests .....	24
2. ANOVA for HR, using adjusted SS for tests.....	24
3. ANOVA for GIDP, using adjusted SS for tests .....	25
4. ANOVA for BB, using adjusted SS for tests .....	27
5. ANOVA for SO, using adjusted SS for tests .....	28
6. ANOVA for SB, using adjusted SS for tests .....	30



## LIST OF FIGURES

1. Main Effects Plot for 2B .....	25
2. Main Effects Plot for HR .....	26
3. Main Effects Plot for GIDP .....	26
4. Main Effects Plot for BB .....	28
5. Main Effects Plot for SO.....	29
6. Main Effects Plot for SB .....	30
7. Percentage of Player by Ethnic Group by Year .....	31

## INTRODUCTION

For well over one hundred years, the game of baseball has repeatedly been called ‘America’s Pastime.’ This simple phrase has firmly etched into the minds of generations of United States citizens that the game is a symbol of a shared cultural heritage. It is an American institution, created and owned by members of this country. Through half of the country’s tumultuous history, the game has not only existed but has functioned as a mirror of social and political change.

However, the game is not static. Resourceful individuals have done with baseball what has been done with countless institutions in the age of modernization. Baseball has gone global. It has spread in almost every conceivable direction across the planet and is even followed in parts of the world where the game is not even played (Klein 2006b).

Yet, the demographics of the game of baseball in the United States are changing. The fan base is not only getting older, but it is altering along racial, ethnic, and gender lines (Klein 2006b). African American presence in the game is in sharp decline while there is an increase in both participants and spectators amongst people of both *Latino* and Asian heritage. Competition with other forms of entertainment seems to have played a part in the decline of the popularity of the sport.

This competition as well as team expansion has also led to a decline in the number of ‘home-grown’ players available to fill the necessary positions on the thirty available MLB teams. MLB players are increasingly coming to the United States from a variety of places such as the Dominican Republic, Cuba, Japan, Venezuela as

well as many others. For instance, between the years 1871 and 1950, there were only fifty-four *Latino* ballplayers in Major League Baseball (Regalado 2002). However in 2005 alone, there were 204 *Latino* ballplayers and this number made up about 25 percent of all MLB players (Klein 2006a).

Riess (1980) argues that “the conventional wisdom that professional baseball was an important alternate source of upward mobility for lower-class youths is inaccurate.” He posits that major league jobs went primarily to middle-class natives and that the myth of upward mobility was for the benefit of immigrants. However, his argument is based around players prior to the 1920s before integration. *Latino* players tend to come from more meager circumstances and the myth of upward mobility is well entrenched (Regalado 1998; Klein 1995).

This project attempts to understand the changing dynamics of the game and its demographics through the lens of cultural transmission. Using the effect of baseball on Latin America, I explore how the decisions that individual players make to get noticed by MLB teams for purposes of upward mobility impact the meanings of their ethnicity and culture. These decisions have not only accounted for changes in racial/ethnic stereotypes, but have actually made lasting impressions on cultural values. Further, I explore how simple game strategies have created a feedback loop that has led to reliance on baseball as a potential escape from poverty and inequalities at home.

## **A VERY BRIEF HISTORY OF MAJOR LEAGUE BASEBALL**

On the first day there was nothing, but on the second day, well, he created baseball. As far as creation stories go, the story of baseball's invention is as varied as any. Though the myth of Abner Doubleday's formulation of the game in Cooperstown in 1839 has been repeatedly debunked, he is still often given credit despite stories of the game being played dating back into the 18<sup>th</sup> century (Rader 2002). However, from Albert Spalding to Alexander Cartwright, the identity of the specific architect has been constantly debated since the game started gaining popularity in the mid-19<sup>th</sup> Century (Rader 2002).

Likely, the game wasn't created through any one individual and was instead derived from similar British games. The modern game can trace its routes back to rule set of the Knickerbockers created in September 1845 by Cartwright (Ivor-Campbell 2002). Although the Knickerbockers game has many similarities to the present one, there were still some major differences. For instance, pitchers would literally pitch the ball in an underhand fashion (Rader 2002). In the following years, the game would begin to organize into a number of amateur and professional leagues run by players.

Though MLB keeps statistics from the defunct National Association that ran from 1871 to 1875, it wasn't until 1876 that modern professional baseball had its first appearance (MLB.com 2007). Over the next quarter century, professional leagues came and went with varying degrees of success, often giving new ideas to the leagues that followed. For instance, the American Association was a successful league for a

decade before it merged with the National League in 1891 bringing with it both ideas and some of the teams that still exist in the NL today. Despite the growth of the NL, Major League Baseball as an organization can trace its true start to the merging of the National League and the upstart American League in 1902.

However, Major League Baseball was still not whole and would not be so until Jackie Robinson broke the color barrier in 1947. The segregation of baseball was a natural extension of the culture of the United States at the time. Segregation was a natural part of life and extended to almost every aspect of life. Tygiel (2002) argues that the formation of the Negro Leagues in the 1920s wasn't just to give African American players a place to play. They were created for a sense of African American empowerment; they were featured at every level of organization. It also kept players ready for eventual integration.

## **LATIN AMERICAN BASEBALL**

Baseball was introduced into Latin America in the mid-19<sup>th</sup> Century (Burgos 2007). Though each country has its own particular way in which the game was established, it was usually through indirect contact with the United States rather than a concentrated effort on part of any organized professional league. The following short histories will focus on two specific countries and how baseball found itself being played there.

### **Cuba**

The appearance of baseball in Cuba was first attributed to American soldiers that were stationed there and who taught the game to the locals in the late 1800s following the Spanish American War (Klein 1997). However, other research tends to attribute the introduction of baseball to Cuban students studying in the United States who brought the game with them on their return to Cuba in the 1860s (Klein 1997; Van Bottenburgh 2001). Baseball became popular quickly and games were often followed by huge celebrations featuring food and expressions of Cuban culture (Carter 2006).

The Cuban League, which was small in numbers of teams but not in popularity, hit its prime in the decades preceding the Cuban Revolution (Carter 2006). All of these teams were located near or in the capital city of Havana and the game tended to be attended mostly by affluent individuals. However, the league and the game were both a point of pride as well as a symbol of Cuban independence.

When Castro, a former ballplayer himself, took over in 1959, he began to change the way that baseball was run in Cuba (Carter 2006). He moved teams and created stadiums to ensure that the sport was available to the public. New leagues were formed and these leagues featured more teams than before. More people signed on to play the game for several reasons. For instance, Cuban ballplayers were often given special privileges not given to others under Castro's regime (Carter 2006). These acts created new rivalries and revitalized the sport. Castro would even participate in games and has often been cited for throwing a great curveball (Wendel 2006).

The changing of the government also had some negative impacts on the sport. The agreements between MLB and the Cuban League were voided and the professional Cuban system was abolished in favor of an amateur league (Carter 2006). In addition, many Cuban players fled or were exiled. However, the new system created ties between the teams and their locations. Cuban baseball was no longer subject to big market teams hiring all of the players as these players generally stayed in their home province throughout their careers. These ties create a sense of community and player/fan interaction that is not seen at the professional level in the United States (Wendel 2006). The collapse of the Soviet Union led to economic problems in Cuba and these issues in addition to the abolishment of the professional leagues led to a decrease in salaries. Due to the lower salaries, some Cuban players still decide to defect to the United States given the disparity in salaries between playing baseball in Cuba and playing in the U.S (Wendel 2006).

## **Dominican Republic**

Baseball can trace its heritage in the Dominican Republic back into the latter half of the nineteenth century when Cuban sugar planters fled the chaos of their civil war (Klein 2006a; Regalado 2002). Cuban immigrants quickly took over the sugarcane industry and many began to organize teams out of their workers and these teams would play against the teams of other refinery managers (Klein 2006a). Often players on these teams would receive extra privileges for participating in these games. These privileges would include such things as time off from the difficult task of cane cutting, a great motivator for players to increase their skill. This story of baseball's roots in the Dominican Republic is famously reflected in the story of San Pedro de Macoris, a city that boasts more MLB players per capita than any other city in the world (Kurlansky 2010).

Four major teams dominated the landscape of professional baseball in the Dominican Republic from 1907 until the 1930s (Klein 2006a). These teams recruited not only local players, but also increasingly added players from abroad. These acquisitions added to the intensity of play as well as the budgets of the teams. Unfortunately, this caused the owners to eventually become bankrupt and put an end to professional baseball in the Dominican Republic for over a decade.

A new league started in 1951 with a new structure and began to attract attention. In 1955, this new league reached a deal with MLB that created the Winter League so that baseball in the Dominican Republic would not compete with professional baseball in the United States (Klein 2006a). More importantly, this deal



began relationships between various MLB teams and Dominican baseball that allowed U.S. players to play there in the winter in exchange for the development of Dominican prospects.

The development of the Dominican academy system started in the 1980s as a way to evaluate and develop talent locally (Klein 2006a). Started by the Blue Jays and the Dodgers, other teams quickly realized that this was an advantage in which they wanted to take part. It was readily apparent that Dominican players were both equally talented and much cheaper than their United States counterparts and the number of Dominican players that reached the Majors quickly grew each year. This was a useful practice for MLB teams trying to operate on a tight budget, but the academies took players away from the Dominican amateur leagues (and by extension, the professional leagues).

For those too young for the academies, they are trained by the *buscón*, an analog for the American baseball scout (Klein 2006a). A *buscón*, a word synonymous with agent or finder, but also colloquially used as a synonym for scam artist, provides for all of the young ballplayers needs from coaching and training to education and food (Fainaru 2001). The *buscón* operates outside the established MLB/Dominican relationship which poses a conundrum for many groups involved in the baseball labor trade.

For some, the *buscónes* are an indispensable part of the business and an agent for empowerment for Dominicans in the system (Fainaru 2001). Others demonize them for taking an unfair amount of compensation out of the players they find and

train. While a typical U.S. agent would receive around 5 percent of a ballplayer's earnings, a *buscón* might earn closer to 50 percent (Fainaru 2001). However, they are more than just an agent. They provide upkeep for these youngsters as well as training and contacts and many young players would never get a chance without them. The corruption of this system, though, has led many to push for an International amateur draft (Fainaru 2001; Klein 2006a). This would have long lasting implications on baseball in general and may well change the game strategy dynamics that this paper shows.

### **United States**

The player most often credited for being the first *Latino* in professional baseball was a Cuban named Esteban Bellán who debuted in 1868 (Regalado 2002; Burgos 2007). However, players from all over Latin America have played in United States professional baseball over the century and a half of its existence. The numbers of *Latino* players have increased dramatically over the years, though, this increase truly started with the integration of baseball in 1947 (Klein 2006a).

In the 19<sup>th</sup> Century, professional baseball in most of the Americas was split along color lines; however, the United States seems to have kept this practice the longest. *Latino* players were judged based on both skin color and the depth of their Spanish heritage and bloodline (Burgos 2007). Players judged to be dark were often excluded from playing in the Major Leagues and instead would play in the Negro

Leagues. However, some players like Alejandro Carrasquel were able to get into the Majors despite a darker skin tone (Wilson 2005).

After the integration of Major League Baseball, there was a rush to explore new avenues of talent. The Negro Leagues were not the only target either. Players from across the globe were signed in an effort to out-scout rival teams. Cuba, Venezuela, Mexico, the Dominican Republic, and others became centers of scouting activity. The paths had opened and the signs pointed to cheaper talent. Today, Major League Baseball boasts a global audience and includes players from six continents and over 30 countries.

## **THE QUESTION OF RACE AND ETHNICITY**

Race and ethnicity are concepts that have often been used interchangeably (Smedley 2007). Though both are categories used to differentiate groups of people from one another, race and ethnicity diverge in the way that these differences are forged. Race and ethnicity are value-laden; they are markers that play a prevalent role in social interactions.

Race is based on biology; it is supposed to be unchanging (Smedley 2007). Evolutionary theory dictates that biological differences are the mechanism by which life is able to thrive. At the species level, different groups of organisms are differentially suited for environments based on shared traits within the species. However, there are sometimes traits within a species that are not shared; however, the differences between these subgroups do not follow the definition of species. These minor differences between sub-species allow the species as a whole to survive with environmental changes or even to morph into a new species given time and environmental pressure. Race is a marker of this biological difference in sub-species. It is an indicator of a separated biological history that allowed a divergence between people to create two very different groups within a single species.

However, social scientists claim that race is a concept without a biological basis in humans. How can this be when we can see the difference so easily? The problem is that this concept that has been so important in the history of the United States is based on one flawed assumption. Race is not a normative, distinct way to group people as it might be in other species (Rensberger 1981). Human variation

exists on a true continuum and traits that might be seen as racial do not have distinct geographical or genetic boundaries.

In the hundreds of years that scientists have studied the concept, there has never been a reliable way to categorize race. The differences overlap as there has not been any lengthy period of time in the past in which groups were isolated enough for discrete differences to evolve. Human variation is a continuous concept in which difference in individuals are greater than differences in groups.

Race is still an extremely important social concept, however. Historical events have ensured that race is ingrained into the minds of people across the globe. It is a distinction that has broad social implications and a concept that is consistently reinforced through interpersonal interactions. Though it may be easy for one to declare that a person is of a specific race, there is little scientific validity in doing so. Racial identity today is far different than it was in the past. The 2000 Census allowed for the option to check multiple race boxes for those of mixed heritage (Farley 2002). Self-identification had been used for years prior to this change, but the interpretation was fairly straightforward and lacking under the old policy of check one box in that it did not allow for mixed race categories.

On the other hand, ethnicity is a dynamic concept that has a variety of interpretations. Barth (1969) approached the problem by stating that ethnicity is based on boundaries; it is a way to demarcate one's group from the other. Others state that ethnicity is a fluid set of learned behaviors that include shared combined beliefs, customs, ideology, language, and traditions; in other words, ethnicity is a

synonym for culture. Levine (1999) defines ethnicity as the “method of classifying people (both self and other) that uses origin (socially constructed) as its primary reference.” He posits that minimalistic definition has an advantage in that it makes it easy for methodological reasons. This definition is the one used for this project.

### **Operationalizing Race and Ethnicity**

Earlier studies on baseball have largely ignored the difficulty of attributing race and ethnic identity to individuals (Phillips 1983; Leonard et al. 1988). For racial identification, these studies have often focused on phenotypic differences. The method of identification was often to look at photographs and decide whether someone belonged to one race or another. Without attempting self-identification, racial grouping in research is problematic and there appears to be no consistent method.

In terms of ethnic identity, studies (Phillips 1983; Leonard et al. 1988) have attempted to solve the problem by placing importance on reliability rather than focusing on the flaws of any specific definition. In this case, reliability is repeatability such that a group of studies will use a similar, if flawed, definition so that these studies can work together as a single body of work. Though many of these studies were asking very different questions, it is still necessary to keep a similar working definition for both purposes of comparison and to contribute to a larger working knowledge base. Thus, the operational definition of the term *Latino* will be similar to the previous studies mentioned above.

A *Latino* player will be defined as any person who traces ancestry to locales that share a common history of Spanish colonization, a common language of Spanish, and a shared heritage (Lapchick & Benedict 1993; Gonzalez 2002). These locales include North, Central, and South American countries such as Mexico, Guatemala, and Argentina, as well as Caribbean countries such as Puerto Rico, the Dominican Republic and Cuba. Though some earlier studies include Brazil, it will be excluded in this study due to the fact that the development of baseball in Brazil comes through Japan rather than the United States and the players of Brazilian baseball are usually of Japanese heritage (Azzoni et al. 2006).

However, for the purposes of this study, there is one modification that must be made since the interest of this project is not in the player's ethnicity, but in the culture in which the game of baseball is learned. Thus *Latino*, for this study, will be used only for people born in a country that meets the factors established above. So, a player who might meet the original definition but was born and learned the game of baseball in the United States will be placed in the U.S. group.

## CULTURE AND BASEBALL

Regalado (2001) argues that American neocolonial expansion brought American values that led to an influx of Latin Americans into the United States in the late 19<sup>th</sup> and early 20<sup>th</sup> centuries. These immigrants were searching for the 'American Dream' of prosperity and hope. However, the racism and ethnocentrism of the day meant that few found it. Instead, they found baseball.

The motivations for playing the game of baseball are many. For some, it is a powerful method of upward mobility. Major League baseball players make more than enough money to help raise entire families out of poverty. For others, baseball is a form of resistance. During the 19<sup>th</sup> Century, baseball games were a place where Cubans could gather and celebrate their own identity rather than the identity imposed by colonialism (Carter 2006).

Perhaps unsurprisingly, many *Latinos* feel that baseball is more than a game. It is a reflection of cultural identity as it encompasses a lot about what is important for the shared parts of Latin American culture. As Regalado (1998: xiv) states,

Baseball was more than a game to them. It was a competition that carried social and economic implications. Baseball was a path out of poverty; it helped to bring distinction to their homelands; it was a means to ease the pain and suffering of kinfolk and compatriots; and it provided a sliver of hope to many younger Latinos who might otherwise have envisioned a dim future. Their determination to succeed in the face of an unwelcoming culture reveals the human spirit of Latin players.

Baseball is a part of the collective identity, but it is also part of an individual identity. *Latino* ballplayers belong to two groups: baseball player and Latin



American. Ingrained as baseball may be into the culture of many Latin American countries, it is even more important to those who play it. Baseball players are a subculture of their own with specific ideologies, rituals, and group dynamics. *Latino* ballplayers belong to this group as well, though their role has not been entirely equal to white ballplayers.

### **Inequalities**

Burgos (2007) argues that foreign-born *Latino* players have economic and cultural disadvantages compared to North American players. This criticism can be extended to players that come from Japan as well. Organized Japanese baseball does not have the same American influence that the academies in Latin America possess. Until recently, bidding wars over players like Daisuke Matsuzaka rarely happened for players from Latin American countries. *Latino* players tend to be more numerous and will often sign for much cheaper than players from other places.

Latin American prospects are often seen as a source of cheap talent that can potentially offset the more expensive signings of players that enter the amateur draft (Burgos 2007). Both *Latino* and Japanese ballplayers are ineligible for the amateur draft due to the official draft rules of the MLB. However, the posting system in place due to an agreement between the MLB and the Japanese leagues keeps the costs associated with Japanese players high.

In addition, Spanish speaking players are often held to a different standard than other foreign born players. Translators are often provided for Japanese players

whereas *Latino* players are often expected to quickly learn English (Burgos 2007).

There seems to be a concerted effort on the part of many teams to provide language classes in their respective academies, though one might again raise the question of how this might further indebted players to their team.

Further, there are suggestions of discrimination in baseball, as there may exist an unequal amount of opportunities for players who have, in essence, the same amount of ability (Leonard et al. 1988). One form of discrimination is in the form of “marginality” where there is a tendency to have marginally talented players be white. Another form of discrimination is in the form of “centrality” where there is a tendency to exclude non-white players from positions of ‘command (Phillips 1983).’

## HYPOTHESES

The *Latino* presence in Major League Baseball is an obviously increasing trend (Wendel 2003). Ballplayers are constantly coming from the Hispanic countries and many are becoming quite successful. The success of these players and their contributions back into their country of origin are visible signs of upward mobility. Many successful players use their earnings to give money back to their home countries. For instance, many successful Dominican players reinvest their monies back into Dominican academies (Klein 2006a). This two-way reinforcement cannot help but be an example of a feedback loop. Reinvestment will have an effect on the youth of the Dominican Republic and the question for young hopefuls becomes not whether to take the path or not, but instead of how to get individually noticed. Extended into the rest of Latin America, it is worth a look to see if a more aggressive strategy to the game might be a sign of an economically and culturally motivated way to become that player that makes it.

For purposes of this study, it is important to establish the existence of different baseball strategies and what a more aggressive strategy might entail. In this case, the hypothesis is that *Latino* ballplayers attempt to get noticed by scouts by trying to make the “big play” offensively. The delineation between big plays and the rest will be depicted as a function of risk and success rate. A play will be defined as big when the result could lead to either a significant decrease or increase in run expectancy as postulated by Tango et al (2007).

The plays that will be measured can occur at several points in the batter-pitcher interaction. The primary result to the interaction would be for the batter to put a ball in play. One of several results can transpire after the ball is put into play: a single, a double, a triple, a homerun, an out, or a GIDP (ground into double play). Singles are not valued as highly by run expectancy and are not the risky behaviors that are being studied in this project. Furthermore, triples have been shown to be more of a product of speed and luck than power or skill. They are not under the control of the batter as much as doubles and homeruns. In each case, Latin American hitters are predicted to have a higher number of these types of hits than others.

**H0: There is no significant difference in terms of doubles, homeruns, or GIDP between Latin American hitters and other hitters.**

**H1: There is a significant difference in terms of doubles, homeruns, or GIDP between Latin American hitters and other hitters.**

In lieu of any ball being put into play, it is also possible for the batter to either reach base or not with a walk or strikeout. A high strikeout, low walk total would imply a propensity to swing more often and is generally a riskier strategy. Thus, a test needs to be done to see if *Latino* players have a significantly lower walk total and a significantly higher strikeout total as a group.

**H0: There is no significant difference in terms of walks or strikeouts between Latin American hitters and other hitters.**

**H1: There is a significant difference in terms of walks or strikeouts between Latin American hitters and other hitters.**

In addition to the interaction between batter and pitcher, risky plays can also occur while the batter is on base. Stolen bases are an expression of that risky type of behavior. A difference in stolen base frequency may show a difference in strategy. A higher number of stolen bases for Latin Americans should be expected than in other groups.

**H0: There is no significant difference in terms of stolen bases between Latin American hitters and other hitters.**

**H1: There is a significant difference in terms of stolen bases between Latin American hitters and other hitters.**

Further, these tests need to be evaluated in terms of correlation between strategy and numbers of Latin American players in MLB at any point in time. If the null hypotheses are rejected and the percentage of players of Latin American origin are increasing, then perhaps an argument can be made that these strategies are a contributing factor to this increase.

## **METHODOLOGY**

This project used a data set called the Lahman Baseball Database using an interface designed by Randy Myers to study the effect of cultural transmission on game strategies (Lahman 2010). This data set was modified to include factors such as age, nationality, and ethnicity in order to create a series of grouping variables. The data set excluded all players who play the position of pitcher. The final sample size was 24,512 and contained any individual who has at least 200 plate appearances since 1901 classified by ethnicity as defined by country of origin. Using country of birth to categorize ethnicity prevents situations where a player may be able to fit into two groups and it also allows the study to focus on the cultural environment that led to the teaching of the game. The cutoff in plate appearances is in place due to research that shows that walk rate does not stabilize until around 200 plate appearances (Slowinski 2010). Microsoft Excel and Minitab was used in the collection, classification, and analysis of the data.

Using a series of analysis of variance (ANOVA) tests, the aforementioned hypotheses were tested on a variety of baseball statistics between grouping factors that include ethnicity and age. These tests allow for a test of the hypotheses about the existence of differing game strategies between Latino and non-Latino ballplayers. The ANOVA will show whether there is a difference in the dependent variable that can be explained by ethnicity. Post-hoc Tukey t-tests will follow the ANOVA test to establish the specific ethnic groupings in which the differences exist.

In addition, a graph was made to plot the percentage of players per year versus ethnicity. This will establish if the different strategies are affecting the population of Major League Baseball. If the percentage of *Latinos* in baseball is increasing over time, then that could be explained by the feedback loop where the high risk strategies causes interest from MLB and a contract offer. This offer, in turn, influences the *Latino* perspective of ballplayers seen as upwardly mobile individuals and also gives money to *Latino* ballplayers which is often reinvested in their country of origin's baseball academies. Finally, these factors reinforce the focus on high risk strategies.

## RESULTS

The first set of analyses focus on balls that are put into play (Technically, a home run is not considered a ball in play for the calculation of some statistics; however, the distinction in this study is categorical). A series of ANOVAs were run for three dependent variables characterized as balls in play with a null hypothesis that there is no difference between any of the groups for any variable. The ANOVAs were designed to show if there is some difference in the variables between the groups; the later Tukey test showed which specific groups are different. Doubles, home runs, and ground in double plays (GIDPs) were used as the response variables in a general linear model against age and ethnicity. Tables 1, 2, and 3 show the rejection of the null hypotheses and that there are significant differences ( $p < 0.05$ ) in both factors for each of the responses.

A series of Tukey-Kramer tests were also run for each of the response variables. These tests were used to establish between which groups in the independent variables the differences lay. There is a clear differentiation between the Hispanic and US groups in terms of all three responses (see Appendices C, D, and E). Figures 1, 2, and 3 are visual representations of the significant difference in means in each of the responses and show the direction of those differences. They are also a good representation of what the results of the Tukey-Kramer test imply. In each case, the Hispanic group had a higher mean than the US group indicating that the Hispanic group is taking a higher risk strategy.



**Table 1**

ANOVA for 2B, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age Group	4	7790.5	7909.8	1977.5	21.31	0.000*
US/His/Oth	2	2649.0	2649.0	1324.5	14.27	0.000*
Error	24505	2274144.1	2274144.1	92.8		
Total	24511	2284583.6				

---

S = 9.63345	R-Sq = 0.46%	R-Sq(adj) = 0.43%
-------------	--------------	-------------------

\*Significant difference at  $P < 0.05$ .**Table 2**

ANOVA for HR, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age Group	4	3883.2	4518.3	1039.6	11.87	0.000*
US/His/Oth	2	5450.5	5450.5	2725.3	31.12	0.000*
Error	24505	2145728.9	2145728.9	87.6		
Total	24511	2155062.6				

---

S = 9.35751	R-Sq = 0.43%	R-Sq(adj) = 0.41%
-------------	--------------	-------------------

\*Significant difference at  $P < 0.05$ .

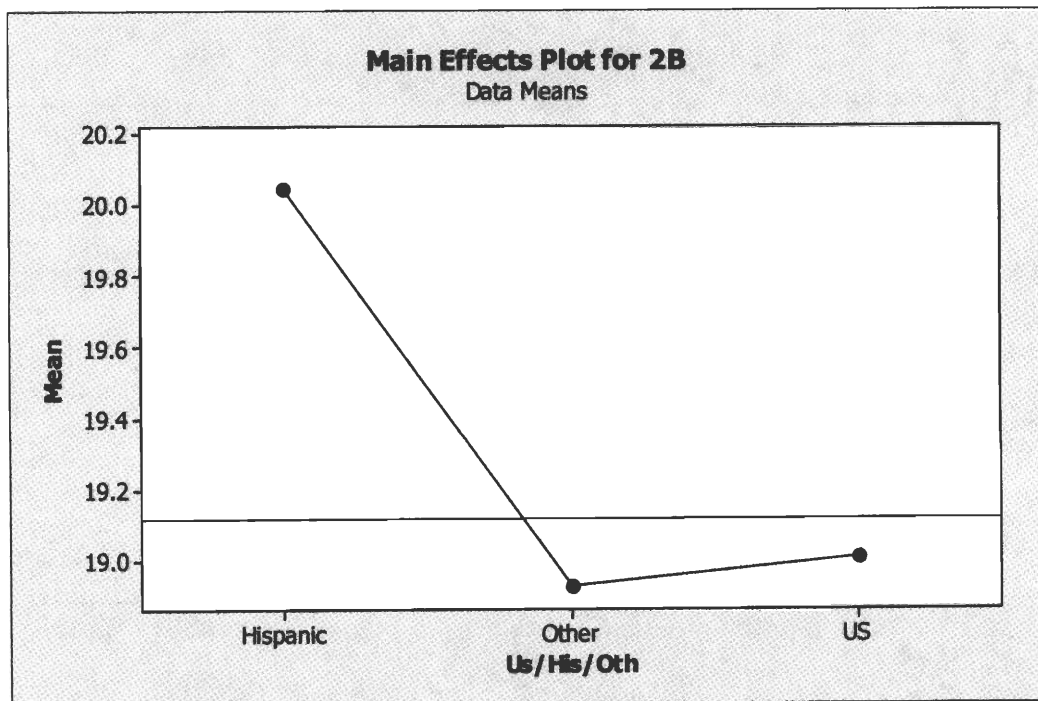
**Table 3**

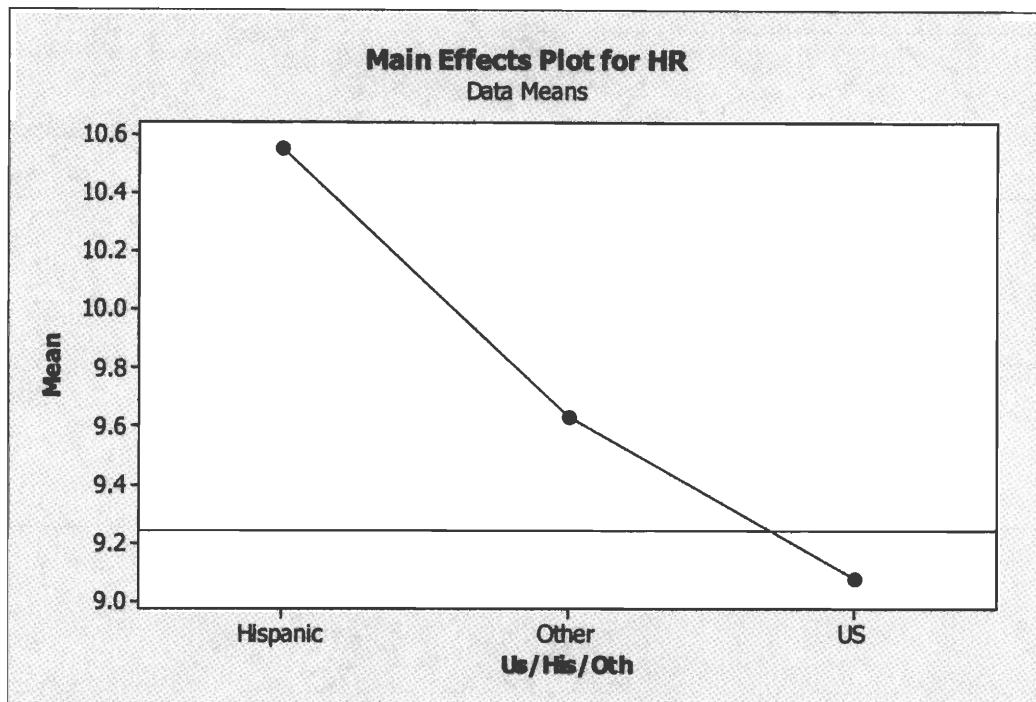
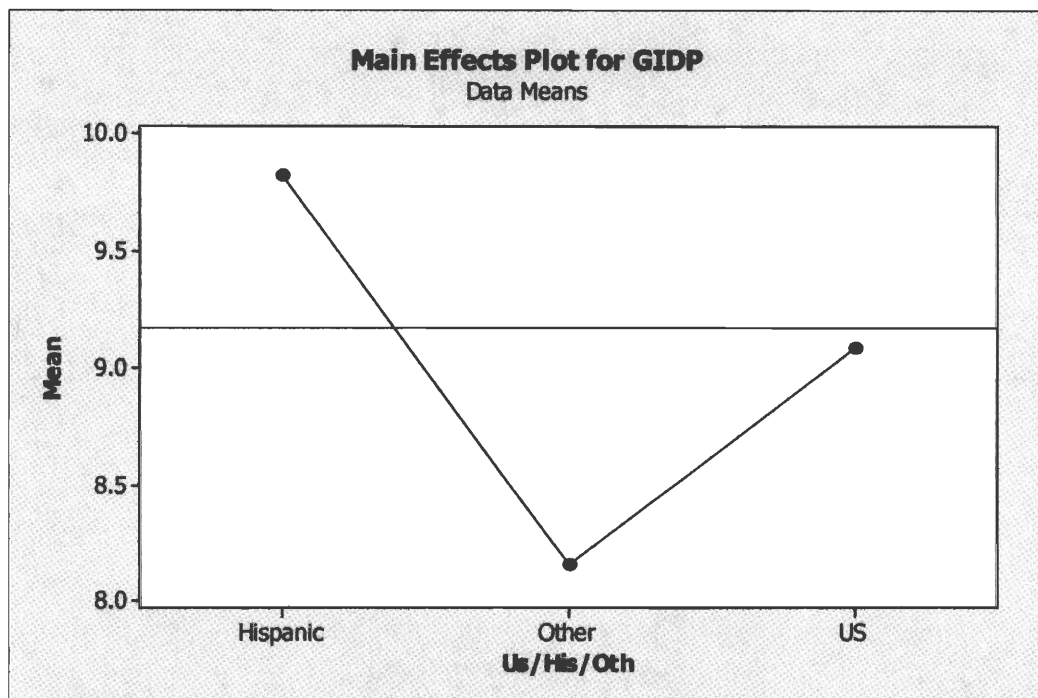
ANOVA for GDP, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age Group	4	1692.09	1838.83	459.71	18.26	0.000*
US/His/Oth	2	1702.27	1702.27	851.14	33.80	0.000*
Error	18671	470122.68	470122.68	25.18		
Total	18677	473517.05				

---

S = 5.01790                      R-Sq = 0.72%                      R-Sq(adj) = 0.68%

\*Significant difference at  $P < 0.05$ .**Figure 1. Main Effects Plot for 2B**

**Figure 2.** Main Effects Plot for HR**Figure 3.** Main Effects Plot for GDP

The results of the next set of ANOVA tests were conclusive and the null hypothesis was rejected at a significance level of 0.05. Table 4 and 5 shows the results of these ANOVAs and that there is significance on the level of ethnicity. The further post hoc Tukey tests (Appendices F and G) show that there is significance difference between the groups of US and *Latino* with *Latinos* showing more strikeouts and less walks than US players. As before, there are a series of main effects plots for the visualization of these differences (see Figure 4 and 5). In this instance, the Hispanic group had a larger mean in terms of SO, but lower in terms of BB.

**Table 4**

ANOVA for BB, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age Group	4	90639	84490	21123	41.52	0.000*
US/His/Oth	2	75058	75058	37529	73.78	0.000*
Error	24505	12465428	12465428	509		
Total	24511	12631125				

---

S = 22.5541	R-Sq = 1.31%	R-Sq(adj) = 1.29%
-------------	--------------	-------------------

\*Significant difference at  $P < 0.05$

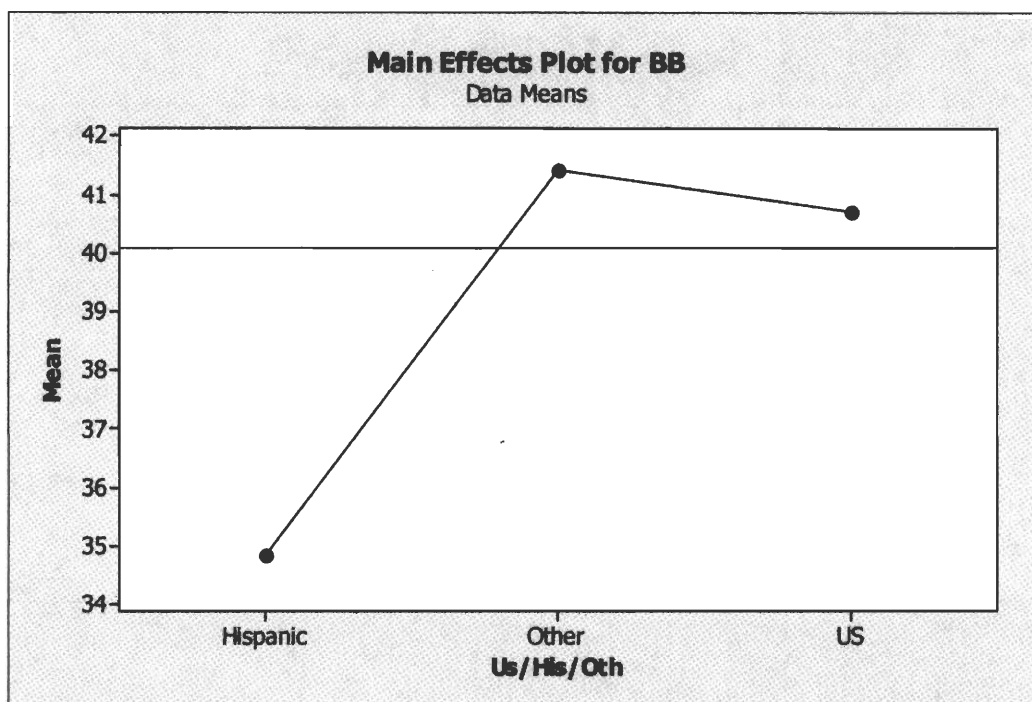
**Table 5**

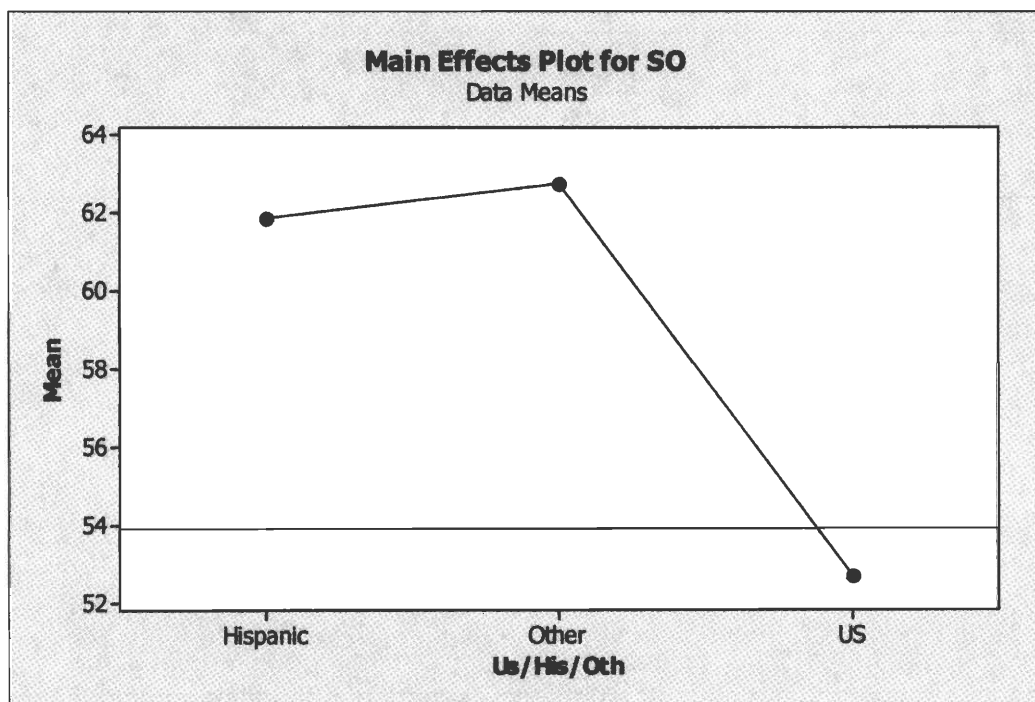
ANOVA for SO, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age Group	4	165963	156328	39082	43.40	0.000*
US/His/Oth	2	217280	217280	108640	120.64	0.000*
Error	22834	20562430	20562430	901		
Total	22840	20945672				

---

S = 30.0086                      R-Sq = 1.83%                      R-Sq(adj) = 1.80%

\*Significant difference at  $P < 0.05$ .**Figure 4. Main Effects Plot for BB**

**Figure 5.** Main Effects Plot for SO

For stolen bases, another ANOVA test was processed. Again, the results of the test were conclusive and the null hypothesis was rejected at a significance level of 0.05. Table 6 shows the results of this ANOVA and that there is significance on both the level of age and the level of ethnicity. Interestingly, the following Tukey test (Appendices H) show that the only significant difference is between the U.S. group and the rest. As before, there is a main effects plot for the visualization of these differences (see Figure 6).

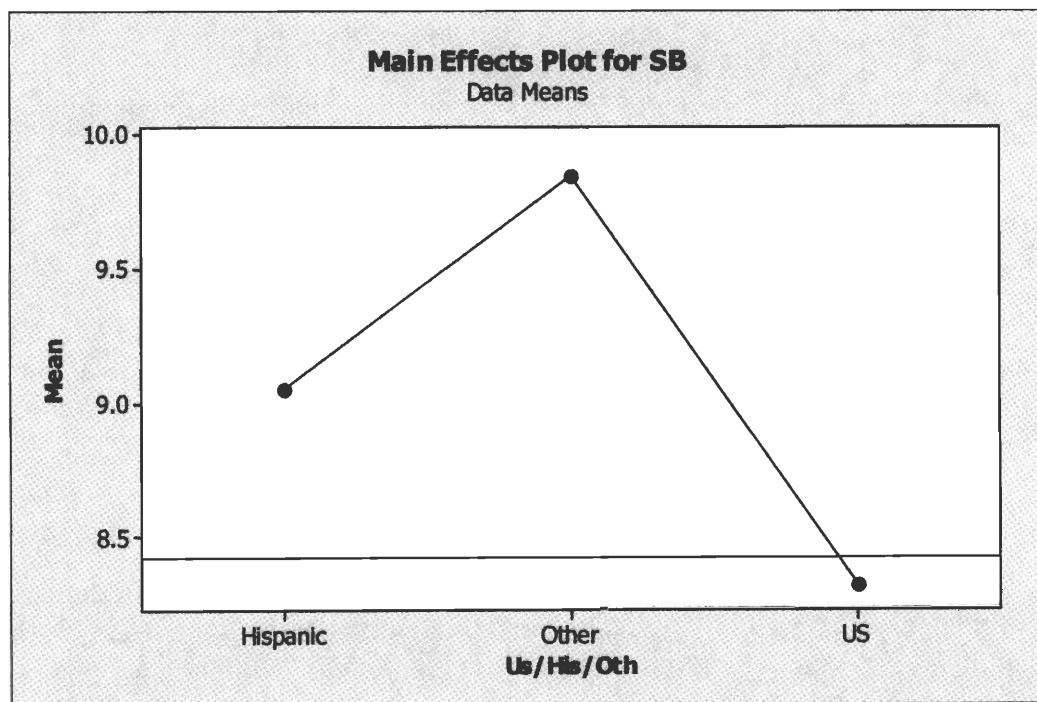
**Table 6**

ANOVA for SB, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age Group	4	30777.4	30628.5	7657.1	67.79	0.000*
US/His/Oth	2	2088.8	2088.8	1044.4	9.25	0.000*
Error	24505	2768130.2	2768130.2	113.0		
Total	24511	2800996.4				

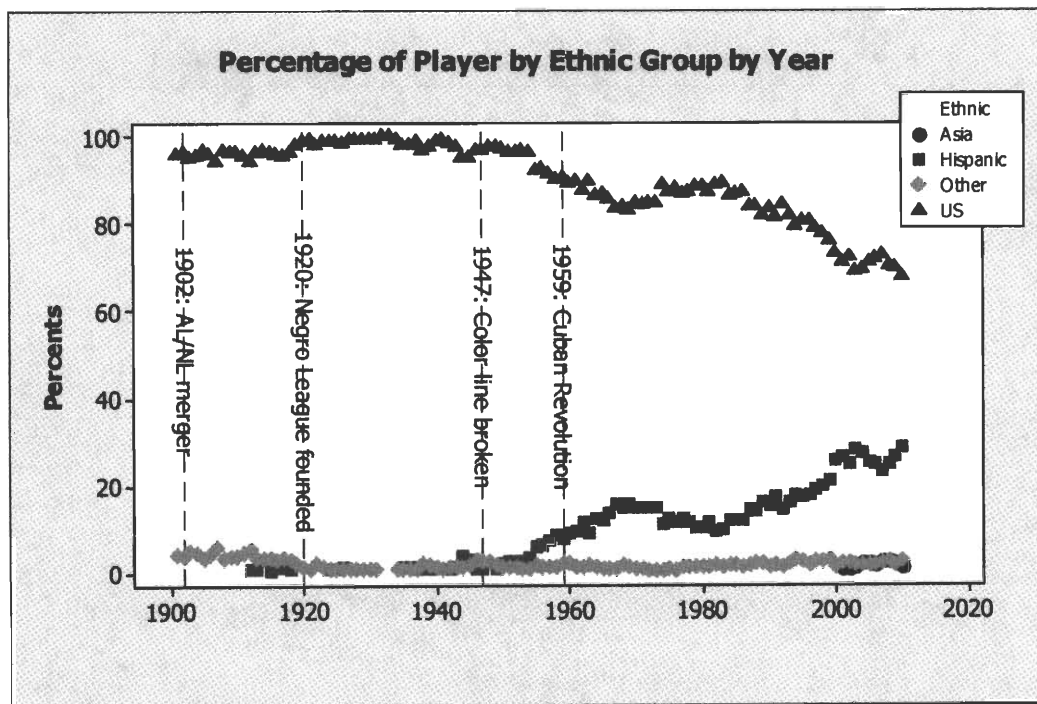
---

S = 10.6284	R-Sq = 1.17%	R-Sq(adj) = 1.15%
-------------	--------------	-------------------

\*Significant difference at  $P < 0.05$ .**Figure 6. Main Effects Plot for SB**

The final question to be answered is whether there is an increase in *Latino* players in Major League Baseball over the past one hundred years. Figure 7 shows a dramatic increase since the mid-1950s. Hispanic players accounted for 3.07% of all players in 1954 and 28.4% of all players in 2010. Likewise, US born players accounted for 96.8% of all players in 1954 and 67.8% of all players in 2010.

**Figure 7.** Percentage of Player by Ethnic Group by Year





## CONCLUSION

The purpose of this thesis was to examine the effect of culturally derived game strategies on the success level of players in the game of baseball. Specifically, I wanted to look at both the influence of how various Latin American cultures taught the game in order to better ensure success of players at the MLB level versus how the game is taught in the United States and Japan. In this model, there is a dichotomy of game strategies between Latin American and United States players in terms of risk with *Latinos* taking a more high risk strategy compared to the United States players. Through this, I hoped to establish a feedback model in how these game strategies perpetuate a cycle of enculturation that further establishes cultural/ethnic identities. The best way to show the establishment of a feedback model is that there is a difference in strategies between *Latino* and non-*Latino* players and by showing that there is an increase in *Latino* players.

The results of this project showed fairly conclusively that there is a major difference in hitting strategies between *Latino* players and US born players. Excepting stolen bases, there was a significant difference between the two groups in all response variables. *Latino* players had, on average, fewer walks, more strikeouts, more GIDPs, more doubles, and more home runs than US players. There does seem to be a significant effect of culture on the results that a player has within the game of baseball.

To show the feedback model, there would have to be an increase in the number of Hispanic players over time. Indeed, there is a significant increase in the

number over time and although there could be many reasons that this would be so; the existence of this fact does not contradict the model. Therefore, there is a very real difference in results between the two groups. The increase in number of *Latino* players over time and the higher number of big plays indicates that there very well could be a difference in game strategy based on culture.

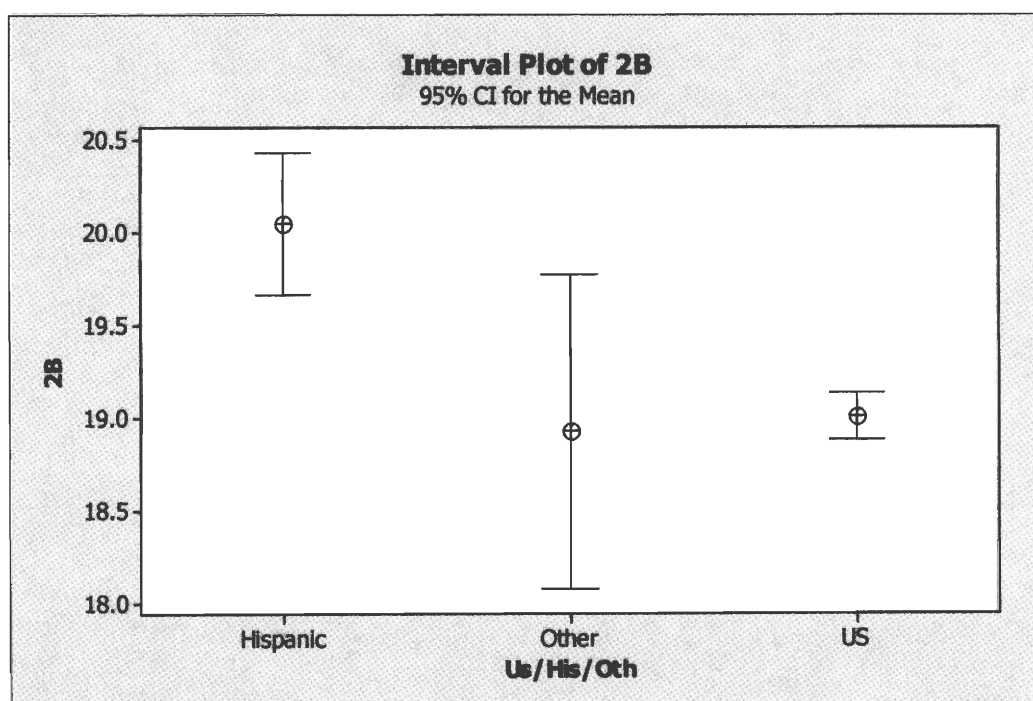
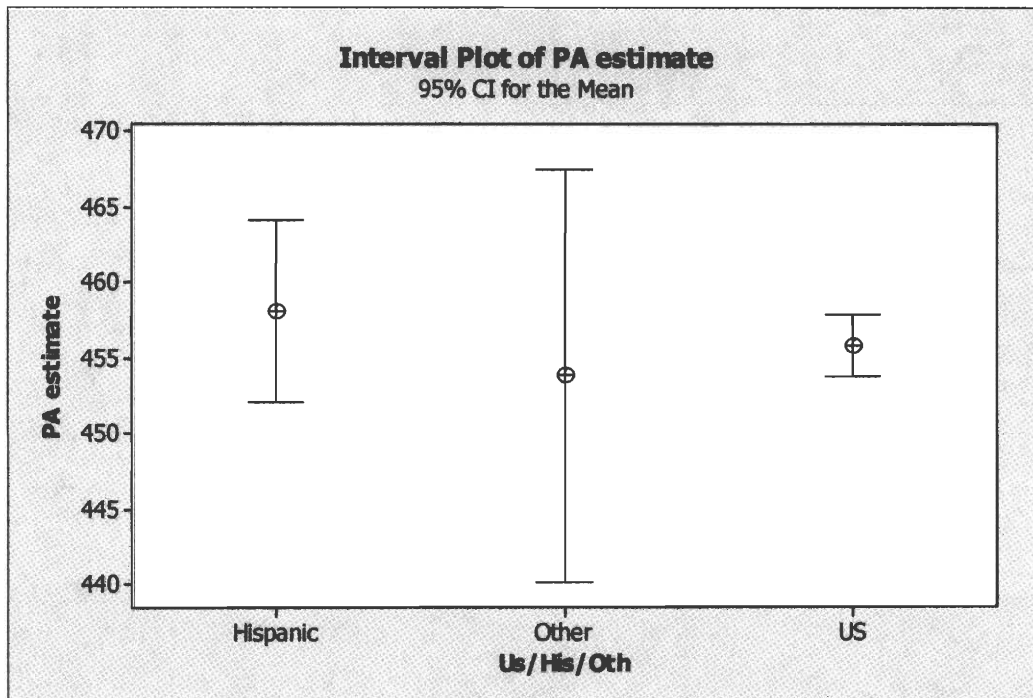
## Appendix A

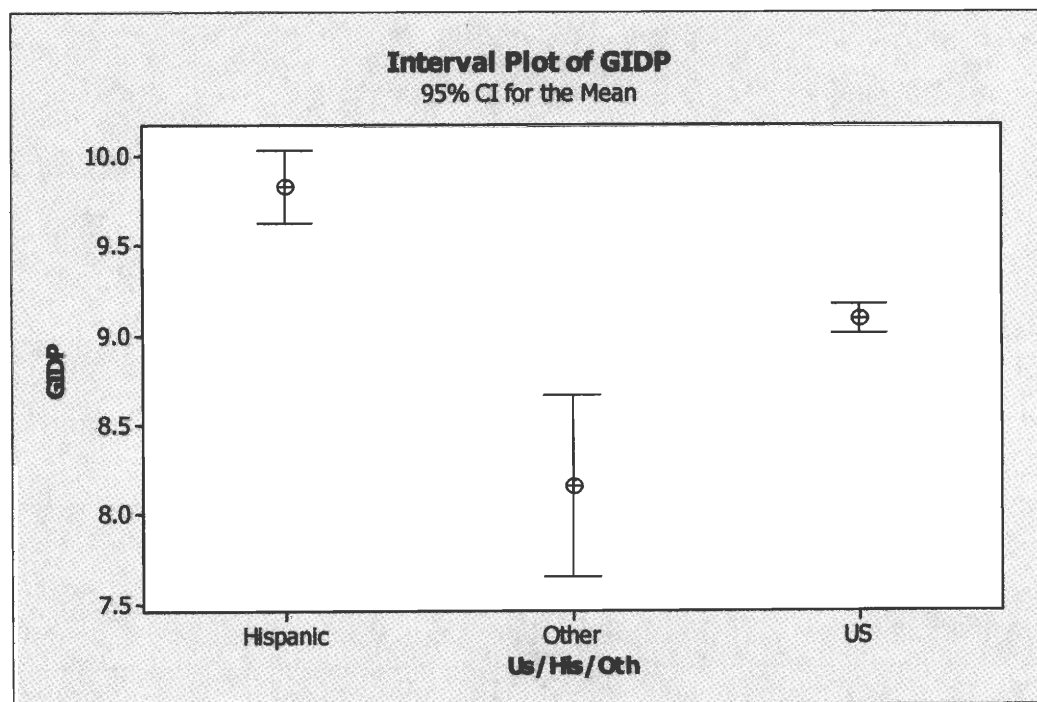
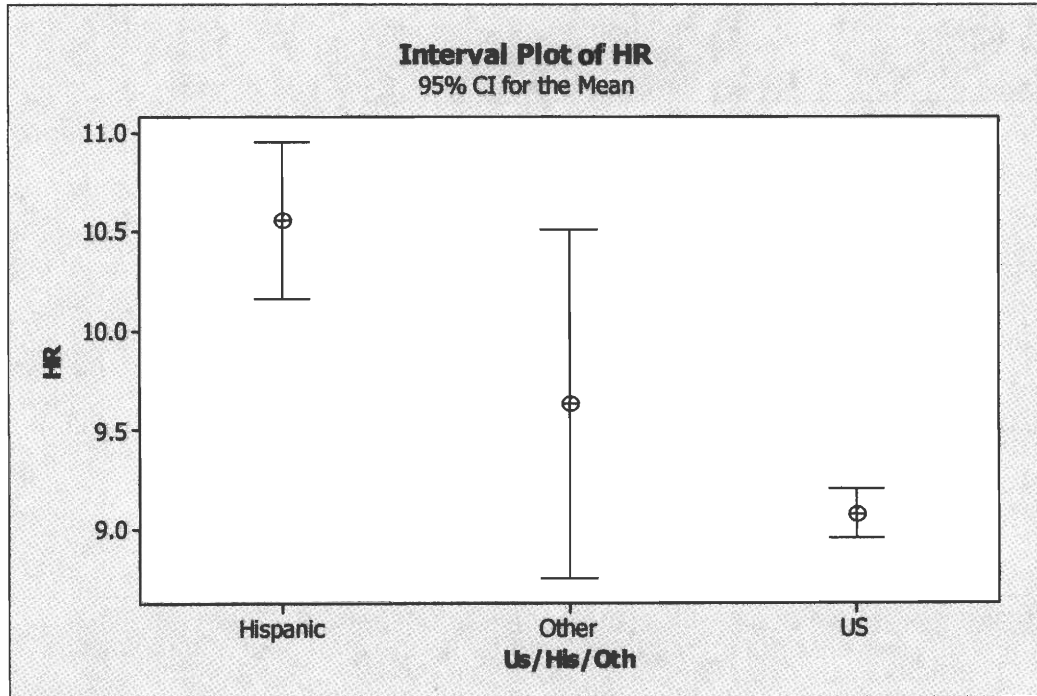
### Player Count by Place of Birth

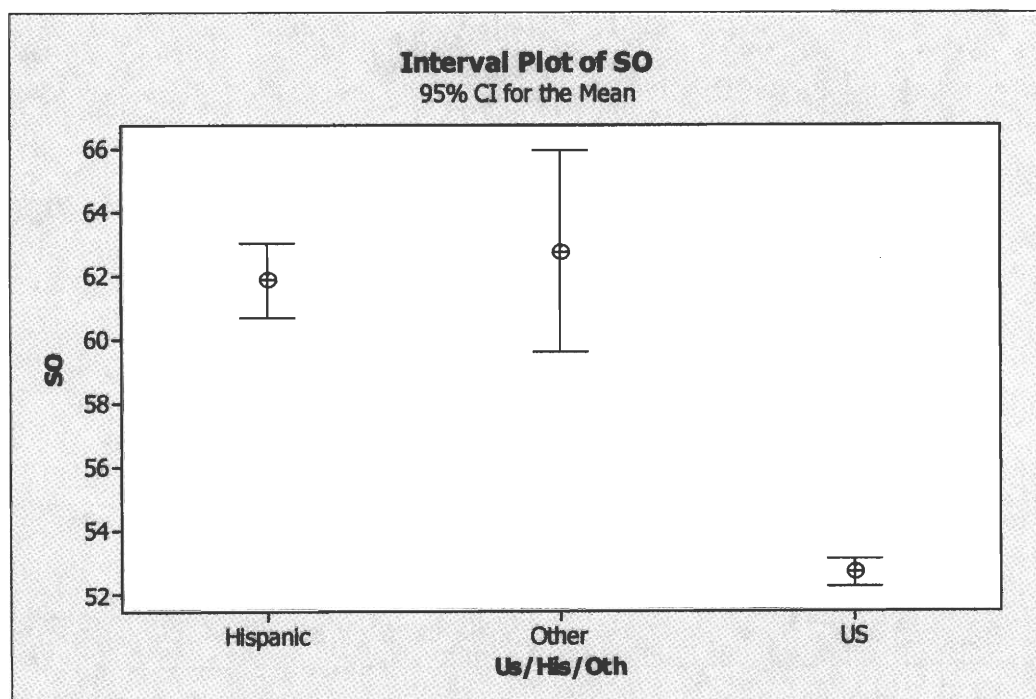
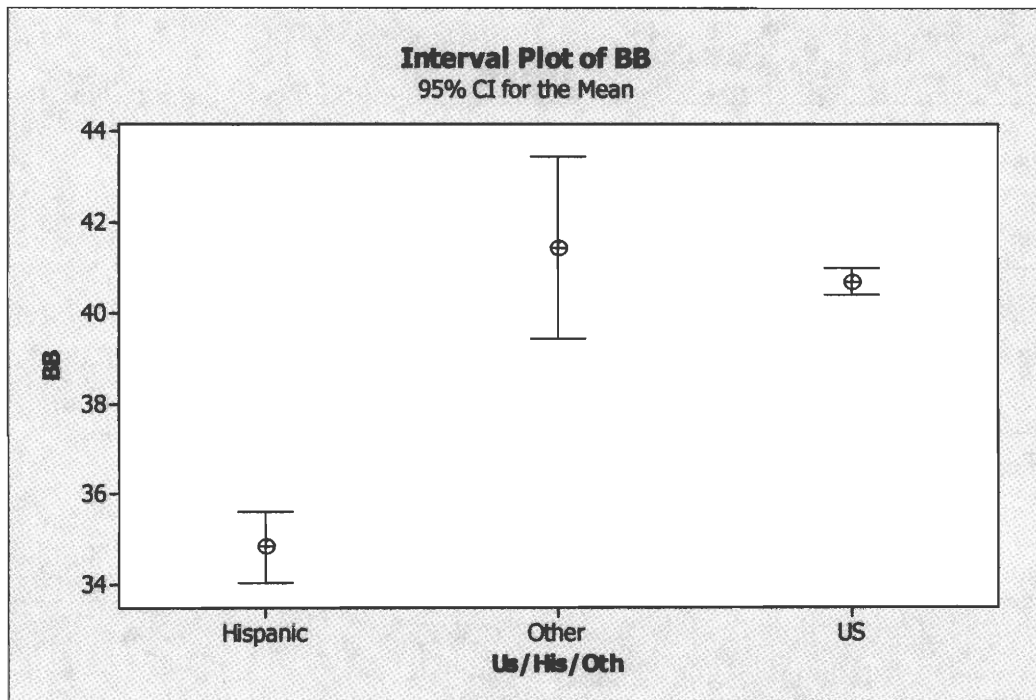
Birth Country	Count	Cumulative Count
American Samoa	4	4
Aruba	1	5
Australia	11	16
Bahamas	7	23
British Honduras	2	25
Canada	171	196
Colombia	35	231
Cuba	312	543
Curacao	19	562
Czechoslovakia	12	574
Dominican Republic	855	1429
England	17	1446
France	6	1452
Germany	10	1462
Honduras	3	1465
Ireland	24	1489
Italy	4	1493
Jamacia	32	1525
Japan	49	1574
Mexico	111	1685
Netherlands Antilles	2	1687
Nicaragua	10	1697
Norway	9	1706
Puerto Rico	684	2390
Panama	140	2530
Russia	8	2538
Scotland	13	2551
South Korea	6	2557
USA	21430	23987
Virgin Islands	30	24017
Venezuela	463	24480
West Germany	20	24500
Wales	12	24512
TOTAL	24512	24512
Total US		21430
Total Hispanic		2615
Total Other		467

## Appendix B

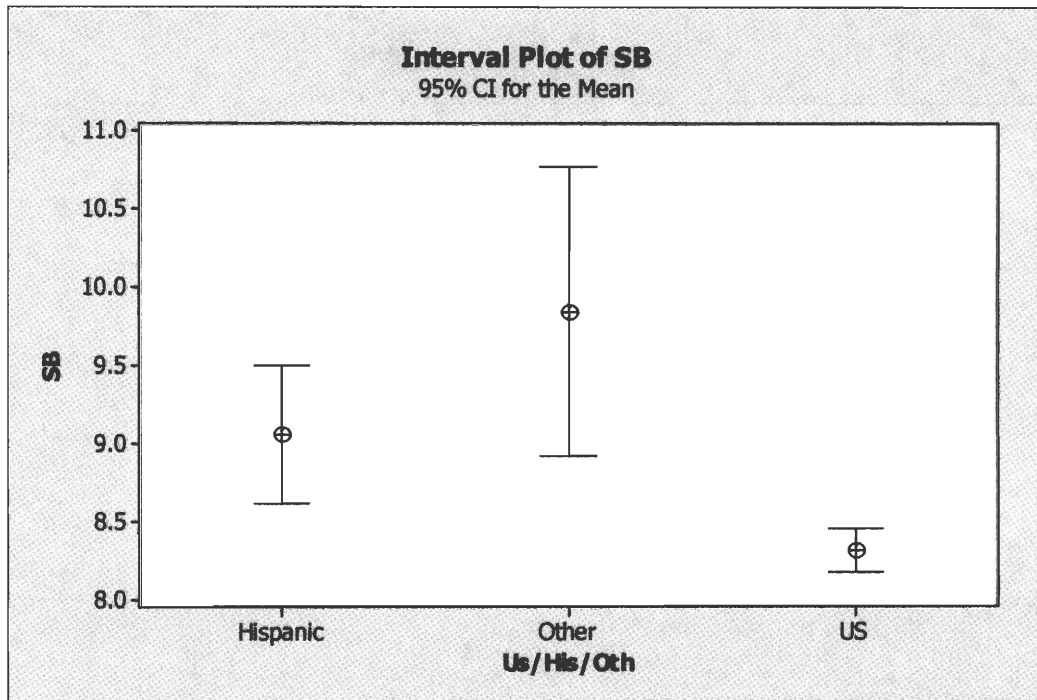
### Plots of Means











## Appendix C

### Tukey Post-Hoc Tests for 2B

**Tukey Simultaneous Tests (2B)**  
**All Pairwise Comparisons among Levels of Ethnicity**

Ethnicity		Difference of Means	Adjusted T-Value	P-Value
<b>Hispanic subtracted from Ethnicity</b>				
Other		-1.149	-2.374	0.0463
US		-1.063	-5.324	0.0000*
<b>Other subtracted from Ethnicity</b>				
US		0.08612	0.1911	0.9801

\*Significant difference at  $P < 0.05$ .

## Appendix D

### Tukey Post-Hoc Tests for HR

**Tukey Simultaneous Tests (HR)**  
**All Pairwise Comparisons among Levels of Ethnicity**

Ethnicity		Difference of Means	Adjusted T-Value	P- Value
<b>Hispanic subtracted from Ethnicity</b>				
Other		-0.997	-2.119	0.0860
US		-1.522	-7.846	0.0000*
<b>Other subtracted from Ethnicity</b>				
US		-0.5253	-1.200	0.4532

\*Significant difference at  $P < 0.05$ .

## Appendix E

### Tukey Post-Hoc Test for GDP

**Tukey Simultaneous Tests (GIDP)**  
**All Pairwise Comparisons among Levels of Ethnicity**

Ethnicity		Difference of Means	Adjusted T-Value	P-Value
<b>Hispanic subtracted from Ethnicity</b>				
Other		-1.719	-5.843	0.0000*
US		-0.778	-7.314	0.0000*
<b>Other subtracted from Ethnicity</b>				
US		0.9404	3.359	0.0023*

\*Significant difference at  $P < 0.05$ .

## Appendix F

### Tukey Post-Hoc Tests for BB



**Tukey Simultaneous Tests (BB)**  
**All Pairwise Comparisons among Levels of Ethnicity**

Ethnicity		Difference of Means	Adjusted t-Value	P-Value
<b>Hispanic subtracted from Ethnicity</b>				
Other		6.217	5.484	0.0000*
US		5.656	12.096	0.0000*
<b>Other subtracted from Ethnicity</b>				
US		-0.5614	-0.5320	0.8556

\*Significant difference at  $P < 0.05$ .

## Appendix G

### Tukey Post-Hoc Tests for SO

**Tukey Simultaneous Tests (SO)**  
**All Pairwise Comparisons among Levels of Ethnicity**

Ethnicity		Difference of Means	Adjusted T-Value	P-Value
<b>Hispanic subtracted from Ethnicity</b>				
Other		1.500	0.93	0.6208
US		-8.901	-14.24	0.0000*
<b>Other subtracted from Ethnicity</b>				
US		-10.40	-6.862	0.0000*

\*Significant difference at  $P < 0.05$ .

## Appendix H

### Tukey Post-Hoc Test for SB

**Tukey Simultaneous Tests (SB)**  
**All Pairwise Comparisons among Levels of Ethnicity**

Ethnicity		Difference of Means	Adjusted T-Value	P- Value
<b>Hispanic subtracted from Ethnicity</b>				
Other		1.0369	1.941	0.1273
US		-0.6285	-2.853	0.0121*
<b>Other subtracted from Ethnicity</b>				
US		-1.665	-3.349	0.0023*

\*Significant difference at  $P < 0.05$ .

## BIBLIOGRAPHY

Adler, Joseph.

2006. *Baseball Hacks*. Cambridge: O'Reilly.

Azzoni, Carlos, Tales Azzoni, and Wayne Patterson

2006. Brazil: Baseball is Popular, and the Players are Japanese!. *In* *Baseball without Borders*. Pp. 196-211. Lincoln: University of Nebraska Press

Barth, Fredrik.

1969. *Ethnic Groups and Boundaries: the Social Organization of Culture Difference*. London: George Allen and Unwin

Baseball Reference.

2007. Players: Search for Place of Birth. Individual Search Data Set  
<<http://www.baseball-reference.com/players.shtml>>

Blanchard, Kendall.

1995. *The Anthropology of Sport: An Introduction*. Westport: Bergin & Garvey.

Burgos Jr., Adrian.

2007. *Playing America's Game: Baseball, Latinos, and the Color Line*. Berkeley: University of California Press.

Carter, Thomas.

2006. Cuba: Community, Fans, and Ballplayers. *In* *Baseball without Borders*. Pp. 147-159. Lincoln: University of Nebraska Press.

Davenport, Clay, Nate Silver, and Keith Woolner.

2007. Statistical Introduction. *In* *Baseball Prospectus 2007*. pp. 1-6. USA: Penguin Group.

Fainaru, Steve.

2001. "The Business of Building Ballplayers." *The Washington Post*. 17 June, 2001: A01. <<http://www.latinamericanstudies.org/sports/dominican-ballplayers.htm>>

Farley, Reynolds.

2002. Racial Identities in 2000: The Response to the Multiple-Race Response Option. *In* *The New Race Question*. Joel Perlmann and Mary Waters, eds. Pp.33-61. New York: Russell Sage Foundation.

Gonzalez, G. Leticia.

2002. "The Stacking of Latinos in Major League Baseball: Does it Matter if a Player is Drafted?" *Journal of Hispanic Higher Education* 1(4):320-328.

Guevara, Arturo J. Marcano and David P. Fidler.

2002. Stealing Lives: The Globalization of Baseball and the Tragic Story of Alexis Quiroz.

Heaphy, Leslie A.

2003. The Negro Leagues: 1869-1960. Jefferson: McFarland & Company, Inc.

Ivor-Campbell, Frederick

2002. The Many Fathers of Baseball: Anglo-Americans and the Early Game. *In The American Game: Baseball and Ethnicity*. Carbondale: Southern Illinois University Press

James, Bill.

2005. "Underestimating the Fog." *The Baseball Research Journal* 33:29-33.

Kedrowski, Ryan.

2005. Analyzing Player Performance in Baseball: A Sabermetric Approach using Data Envelopment Analysis. Master's Thesis. Cambridge University: Judge Business School.

Keri, Jonah.

2006. What's the Matter with RBI?...and Other Traditional Statistics. *In Why Everything You Know About the Game is Wrong: Baseball Between the Numbers*. Jonah Keri, ed. Pp. 1-13. New York: Basic Books

Kish, Leslie.

1978. "Chance, Statistics, and Statisticians." *Journal of the American Statistical Association* 73(361): 1-7.

Klein, Alan.

1995. "Culture, Politics, and Baseball in the Dominican Republic." *Latin American Perspectives* 22(3): 111-130.

1997. Baseball on the Border: A Tale of Two Laredos. Princeton: Princeton University Press.

2006a. Dominican Republic: Forging an International Industry. *In Baseball without Borders*. Pp. 117-135. Lincoln: University of Nebraska Press.

2006b. *Growing the Game: The Globalization of Major League Baseball*. New Haven: Yale University Press.

Kurlansky, Mark.

2010. *The Eastern Stars: How Baseball Changed the Dominican Town of San Pedro de Macoris*. London: Penguin Group.

Lahman, Sean.

2010. Lahman Baseball Database.

<<http://www.myerscomputerdesigns.com/BaseballApplication.html>>

Lapchick, R. E. and J. R. Bendict.

1993. 1993 Racial Report Card. Boston: Northeastern University, Center for the Study of Sport in Society.

Leonard, Wilbert M., J. Pine, and C. Rice.

1988. "Performance Characteristics of White, Black, and Hispanic Major League Baseball Players: 1955-1984." *Journal of Sport and Social Issues* 12(1): 31-43.

Levine, Hal B.

1999. "Reconstructing Ethnicity." *The Journal of the Royal Anthropological Institute* 5(2): 165-180.

Lewis, Michael.

2003. *Moneyball*. New York: W. W. Norton & Company.

MLB.com.

2001. Major League Baseball History. 12 Dec. 2007.

<<http://mlb.mlb.com/mlb/history/index.jsp>>

Perry, Dayn.

2006. Extra Innings: Can Stats and Scouts Get Along. *In Why Everything You Know About the Game is Wrong: Baseball Between the Numbers*. Jonah Keri, ed. Pp. 369-371. New York: Basic Books

Phillips, John C.

1983. Race and Career Opportunities in Major League Baseball: 1960. *Journal of Sport and Social Issues* 7(1):1-17.



Rader, Benjamin G.

2002. *Baseball: A History of America's Game*. Urbana; University of Illinois Press.

Regalado, Samuel O.

1998. *Viva Baseball! Latin Major Leaguers and Their Special Hunger*. Urbana: University of Illinois Press.

2001. Sammy Sosa Meets Horatio Alger: Latin Ballplayers and the American Success Myth. *In* *Baseball and the American Dream: Race, Class Gender, and the National Pastime*. Robert Elias, ed. Pp. 71-75. Armonk: M.E. Sharpe.

2002. The Latin Quarter in the Major Leagues: Adjustment and Achievement. *In* *The American Game: Baseball and Ethnicity*. Carbondale: Southern Illinois University Press

Riess, Steven A.

1980. "Professional Baseball and Social Mobility" *Journal of Interdisciplinary History* 11(2): 235-250.

Rensberger, Boyce.

1981. Racial Odyssey. *Science Digest*. January/February 1981.

Rodriguez, Clara.

2000. *Changing Race: Latinos, the Census, and the History of Ethnicity in the United States*. New York: New York University Press.

Sanocki, Thomas.

2001. *Student Friendly Statistics*. New Jersey: Prentice Hall.

Slowinski, Steve.

2010. "Sample Size | FanGraphs Sabermetrics Library." *Baseball Statistics and Analysis | FanGraphs Baseball*. 18 Feb. 2010. Web. 06 Dec. 2011. <<http://www.fangraphs.com/library/index.php/principles/sample-size/>>.

Smedley, Audrey.

2007. *Race in North American: Origin and Evolution of a Worldview*. Boulder: Westview Press.

Tango, Tom, Lichtman, Mitchel, and Andrew Dolphin.

2007. *The Book: Playing the Percentages in Baseball*. Washington, D.C.: Potomac Books.

Tomlinson, Gerald, ed.

2000. *How to Do Baseball Research*. Cleveland: Society for American Baseball Research.

Tygiel, Jules

2002. *Unreconciled Strivings: Baseball in Jim Crow America*. In *The American Game: Baseball and Ethnicity*. Carbondale: Southern Illinois University Press.

Van Bottenburg, Maarten.

2001[1994]. *Global Games*. Beverley Jackson, trans. Urbana: University of Illinois Press.

Wendel, Tim.

2003. *The New Face of Baseball*. New York: Rayo.

2006. *Cuba: Behind the Curtain*. In *Baseball without Borders*. Pp. 136-146. Lincoln: University of Nebraska Press.

Wilson, Nick C.

2005. *Early Latino Ballplayers in the United States: Major, Minor and Negro Leagues, 1901-1949*. Jefferson: McFarland & Company, Inc.